(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum Internationales Büro





(43) Internationales Veröffentlichungsdatum 23. Oktober 2003 (23.10.2003)

(10) Internationale Veröffentlichungsnummer WO 03/087386 A2

(51) Internationale Patentklassifikation7: C12N 9/10, 1/15, 15/63

C12P 13/12,

(74) Anwalt: SCHWEIGER Georg; Reitstötter, Kinzebach & Partner (GbR), Sternwartstrasse 4, 81679 München (DE).

(21) Internationales Aktenzeichen:

PCT/EP03/04010

(22) Internationales Anmeldedatum:

16. April 2003 (16.04.2003)

(25) Einreichungssprache:

Deutsch

(26) Veröffentlichungssprache:

Deutsch

(30) Angaben zur Priorität: 102 17 058.4 17. April 2002 (17.04.2002) DE

- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; 67056 Ludwigschafen (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): KRÖGER, Burkhard [DE/DE]; Im Waldhof 1, 67117 Limburgerhof (DE). ZELDER, Oskar [DE/DE]; Franz-Stützel-Str. 8, 67346 Speyer (DE). KLOPPROGGE, Corinna [DE/DE]; Diemersteinstrasse 3, 67065 Ludwigshafen (DE). SCHRÖDER, Hartwig [DE/DE]; Benzstr. 4, 69226 Nussloch (DE). HÄFNER, Stefan [DE/DE]; Luitpoldstr. 11, 67063 Ludwigshafen (DE).

- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TR, TT, TZ, UA, UG,
- (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

US, UZ, VC, VN, YU, ZA, ZM, ZW.

Veröffentlicht:

ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: METHODS FOR PRODUCING SULPHUROUS FINE CHEMICALS

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG VON SCHWEFELHALTIGEN FEINCHEMIKALIEN

(57) Abstract: The invention relates to methods for producing sulphurous fine chemicals, in particular L-methionine, by fermentation, using bacteria, in which a nucleotide sequence that codes for a methionine synthase (metF) gene is expressed.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien. insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methionin-Synthase (metF)-Gen kodierende Nukleotidsequenzen exprimiert wird.



Verfahren zur Herstellung von schwefelhaltigen Feinchemikalien Beschreibung

Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methionin-Synthase (metH)-Gen kodierende Nukleotidsequenz exprimiert wird.

Stand der Technik

Schwefelhaltige Feinchemikalien, wie zum Beispiel Methionin, Homocystein, S-Adenosyl-Methionin, Glutathion, Cystein, Biotin, Thiamin, Liponsäure werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik- und pharmazeutischen Industrie. Diese Substanzen, die zusammen als "schwefelhaltige Feinchemikalien" bezeichnet werden, umfassen organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Vitamine und Cofaktoren. Ihre Produktion erfolgt am zweckmäßigsten im Großmaßstab mittels Anzucht von Bakterien, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezemieren. Für diesen Zweck besonders geeignete Organismen sind coryneforme Bakterien, gram-positive nicht-pathogene Bakterien.

Es ist bekannt, dass Aminosäuren durch Fermentation von Stämmen coryneformer Bakterien, insbesondere Corynebacterium glutamicum, hergestellt werden. Wegen der großen Bedeutung wird ständig an der Verbesserung der Herstellverfahren gearbeitet. Verfahrensverbesserungen können fermentationstechnische Maßnahmen, wie zum Beispiel Rührung und Versorgung mit Sauerstoff, oder die Zusammensetzung der Nährmedien, wie zum Beispiel die Zuckerkonzentration während der Fermentation, oder die Aufarbeitung zum Produkt, beispielsweise durch Ionenaustauschchromatographie, oder die intrinsischen Leistungseigenschaften des Mikroorganismus selbst betreffen.

Über Stammselektion sind eine Reihe von Mutantenstämmen entwickelt worden, die ein Sortiment wünschenswerter Verbindungen aus der Reihe der schwefelhaltigen Feinchemikalien produzieren. Zur Verbesserung der Leistungseigenschaften dieser Mikroorganismen hinsichtlich der Produktion eines bestimmten Moleküls werden Methoden der Mutagenese, Selektion und Mutantenauswahl angewendet. Dies ist jedoch ein zeitaufwendiges und schwieriges Verfahren. Auf

10

15

25

30

diese Weise erhält man z.B. Stämme, die resistent gegen Antimetabolite oder Hemmstoffe, wie z. B. die Methionin-Analoga α-Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethylhomocystein, 2-Amino-5-heprenoitsäure, Seleno-Methionin, Methioninsulfoximin, Methoxin, 1-Aminocyclopentan-Carboxylsäure oder auxotroph für regulatorisch bedeutsame Metabolite sind und schwefelhaltige Feinchemikalien, wie z. B. L-Methlonin, produzieren.

Seit einigen Jahren werden ebenfalls Methoden der rekombinanten DNA-Technik zur Stammverbesserung von L-Aminosäure produzierender Stämme von Corynebacterium eingesetzt, indem man einzelne Aminosäure-Blosynthesegene amplifiziert und die Auswirkung auf die Aminosäure-Produktion untersucht.

Die WO-A-02/10209 beschreibt ein Verfahren zur fermentativen Herstellung von L-Methionin unter Verwendung L-Methionin produzierender coryneformer Bakterien, worin wenigstens das metH-Gen überexprimiert ist, sowie die kodierende metH-Sequenz aus C. glutamicum ATCC 13032.

Kurze Beschreibung der Erfindung

Der Erfindung lag die Aufgabe zugrunde, ein neues Verfahren zur verbesserten fermentativen 20 Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methlonin, bereitzustellen.

Gelöst wird obige Aufgabe durch Bereitstellung eines Verfahrens zur fermentativen Herstellung einer schwefelhaltigen Feinchemikalie, umfassend die Expression einer heterologen Nukleotidsequenz, welche für ein Protein mit metH-Aktivität kodiert, in einem coryneformen Bakterium.

Ein erster Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen Feinchemikalie, welches folgende Schritte umfasst:

- Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit Methionin-Synthase (metH) –Aktivität kodiert;
- Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und

isolieren der schwefelhaltigen Feinchemikalie, welche vorzugsweise L-Methionin umfasst.

Vorzugsweise besitzt obige heterologe metH-kodierende Nukleotidsequenz zur metH-kodierenden Sequenz aus Corynebacterium glutamicum ATCC 13032 eine Sequenzhomologie vom weniger als 70% aufweist. Die metH-kodierende Sequenz ist vorzugsweise aus einem der folgenden Organismen von Liste I abgeleitet:

<u>Liste I</u>

10

Streptomyces coelicolor Anabaena sp. ATCC 27892 Synechocystis sp. ATCC 27184 Prochlorococcus marinus Thermus thermophilus Bacillus halodurans Bacillus stearothermophilus ATCC 27634 Bacillus stearothermophilus ATCC 21591 Bacillus stearothermophilus Vibrio cholerae ATCC 39315 Sinorhizobium meliloti ATCC 4399 Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi Pseudomonas fluorescens Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Synechocystis sp. ATCC 27184 Prochlorococcus marinus PCC 7118 Thermus thermophilus ATCC 27634 Bacillus halodurans ATCC 21591 Bacillus stearothermophilus ATCC 12980 Vibrio cholerae ATCC 39315 Sinorhizobium meliloti ATCC 4399 Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi ATCC 12839 Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Prochlorococcus marinus Thermus thermophilus ATCC 27634 Bacillus halodurans Bacillus stearothermophilus ATCC 12980 Vibrio cholerae ATCC 39315 Sinorhizobium meliloti ATCC 4399 Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Thermus thermophilus Bacillus halodurans Bacillus stearothermophilus Vibrio cholerae Sinorhizobium meliloti Escherichia coli K12 Salmonella typhimurium Pseudomonas fluorescens Pseudomonas aeruginosa Nitrosomonas europeae ATCC 27634 ATCC 21591 ATCC 12980 ATCC 39315 ATCC 4399 Escherichia coli K12 ATCC 55151 ATCC 15277 Salmonella typhi ATCC 12839 ATCC 13525 ATCC 17933 Nitrosomonas europeae ATCC 19718
Bacillus halodurans Bacillus stearothermophilus Vibrio cholerae ATCC 39315 Sinorhizobium meliloti ATCC 4399 Escherichia coli K12 Salmonella typhimurium ATCC 15277 Salmonella typhi Pseudomonas fluorescens Pseudomonas aeruginosa Nitrosomonas europeae ATCC 19718
Bacillus halodurans Bacillus stearothermophilus Vibrio cholerae Sinorhizobium meliloti Escherichia coli K12 Salmonella typhimurium ATCC 12839 Pseudomonas fluorescens Pseudomonas aeruginosa Nitrosomonas europeae ATCC 21591 ATCC 12980 ATCC 39315 ATCC 4399 ATCC 55151 ATCC 15277 Salmonella typhi ATCC 12839 ATCC 13525 ATCC 17933 ATCC 17933
Bacillus stearothermophilus ATCC 12980 Vibrio cholerae ATCC 39315 Sinorhizobium meliloti ATCC 4399 Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi ATCC 12839 Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Vibrio cholerae ATCC 39315 Sinorhizobium meliloti ATCC 4399 Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi ATCC 12839 Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi ATCC 12839 Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Escherichia coli K12 ATCC55151 Salmonella typhimurium ATCC 15277 Salmonella typhi ATCC 12839 Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Salmonella typhi ATCC 12839 Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Pseudomonas fluorescens ATCC 13525 Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Pseudomonas aeruginosa ATCC 17933 Nitrosomonas europeae ATCC 19718
Bordetella pertussis ATCC 9797
Clorobium tepidum ATCC 49652
Deinococcus radiodurans ATCC 13939
Clostridium acetobutylicum ATCC 824
Caulobacter crescentus ATCC 19089
Homo sapiens
Vibrio fischeri ATCC 33715
Agrobacterium tumefaciens str. C58 (Cereon) ATCC 33970
Ralstonia solanacearum ATCC 25237

ATCC: American Type Culture Collection, Rockville, MD, USA

PCC: Pasteur Culture Collection of Cyanobacteria. Paris Frankreich

Die erfindungsgemäß eingesetzte metH-kodierende Sequenz umfasst vorzugsweise eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35,

37, 39, 41, 43, 45, 47, 49 und 51 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metH-Aktivität kodiert.

Die erfindungsgemäß eingesetzte metH-kodierende Sequenz kodiert außerdem vorzugsweise für ein Protein mit metH-Aktivität, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 und 52 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metH-Aktivität steht, umfasst.

Die kodierende metH-Sequenz ist vorzugsweise eine in coryneformen Bakterlen replizierbare oder eine stabil in das Chromosom intregrierte DNA oder eine RNA.

Gemäß einer bevorzugten Ausführungsform wird das erfindungsgemäße Verfahren durchgeführt, indem man

15

- einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens
 eine Kopie der kodierenden metH-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
- b) einen Stamm einsetzt, in dem die kodierende metH-Sequenz in das Chromosom des Bakteriums integriert wurde.

20

Es ist weiterhin bevorzugt, die kodierende metH-Sequenz für die Fermentation zu überexprimieren.

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie oder eines damit assoziierten Biosynthese- oder sonstigen Stoffwechselweges, verstärkt ist; und / oder

in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

30

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie durch Stoffwechselmetabolite in seiner Aktivität nicht in unerwünschter Weise beeinflusst wird.

Gemäß einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens werden deshalb coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

5 a) dem für eine Aspartatkinase kodierenden Gen lysC. b) dem für eine Aspartat-Semialdehyd-Dehydrogenase kodierenden Gen asd dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen gap, c) d) dem für die 3-Phosphoglycerat Kinase kodierenden Gen pgk, 10 e) dem für die Pyruvat Carboxylase kodierenden Gen pvc. f) dem für die Triosephosphat Isomerase kodierenden Gen tpi. g) dem für die Homoserin O-Acetyltransferase kodierenden Gen metA. dem für die Cystathionin-gamma-Synthase kodierenden Gen metB, h) i) dem für die Cystathionin-gamma-Lyase kodierenden Gen metC, 15 j) dem für die Serin-Hydroxymethyltransferase kodierenden Gen glyA, k) dem für die O-Acetylhomoserin-Sulfhydrylase kodierenden Gen metY, I) dem für die Methylen-Tetrahydrofolat-Reduktase kodierenden Gen, metF m) dem für die Phosphoserin-Aminotransferase kodierenden Gen serC n) dem für die Phosphoserin-Phosphatase kodierenden Gen serB, 20 0) dem für die Serine Acetyl-Transferase kodierenden Gen cysE, p) dem für die Homoserin-Dehydrogenase kodierenden Gen hom. überexprimiert ist.

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene ausgewählt unter Genen der oben genannten Gruppe a) bis p) mutiert ist, insbesondere so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden und dass insbesondere die erfindungsgemäße Produktion der Feinchemikalie nicht beeinträchtigt wird. Durch die Mutation kann das Protein auch eine höhere Aktivität (Sunstratumsatz) und/oder Sunstratspezifität besitzen und damit die Produktion der gewünschten Feinchemikalie fördem.

25

30

WO 03/087386

PCT/EP03/04010

6

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- q) dem für die Homoserine-Kinase kodierenden Gen thrB,
- r) dem für die Threonin Dehydratase kodierenden Gen ilvA,
- s) dem für die Threonin Synthase kodierenden Gen thrC
 - t) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh
 - u) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck,
 - v) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgi,
 - w) dem für die Pyruvat-Oxidase kodierenden Gen poxB,
 - x) dem für die Dihydrodipicolinat Synthase kodiemden Gen dapA,
 - y) dem für die Dihydrodipicolinat Reduktase kodiernden Gen dapB; oder
 - z) dem für die Diaminopicolinat Decarboxylase kodiernden Gen lysA

abschwächt ist, insbesondere durch Verringerung der Expressionsrate des korrespondierenden Gens, oder durch Expression eines Proteins mit geringerer Aktivität (Substratumsatz).

15

10

5

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene der obigen Gruppen q) bis z) mutiert ist, so dass die enzymatische Aktivität des korrespondierenden Proteins teilwelse oder vollständig veringert wird.

20

Vorzugsweise werden in dem erfindungsgemäßen Verfahren Mikroorganismen der Art Corynebacterium glutamicum eingesetzt.

25

In einer weiteren Ausführungsform des Verfahrens werden solche Mikroorganismen eingesetzt, die Resistenz gegenüber wenigstens einen Methionin-Biosynthesehemmer aufweisen. Solche Hemmer sind, ohne darauf beschränkt zu sein, Methionin-Analoga, wie α-Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethylhomocystein, 2-Amino-5-heprenoitsäure, Seleno-Methionin, Methioninsulfoximin, Methoxin, und 1-Aminocyclopentan-Carboxylsäure.

30

Ein weiterer Gegenstand der Erfindung betrifft ein Verfahren zur Herstellung eines L-Methioninhaltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

7

- Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
- b) Entfermung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
- Entfernung der w\u00e4hrend der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
- Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das
 Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

Gegenstand der Erfindung sind ebenfalls die erstmalig aus obigen Mikroorganismen isolierten kodierenden metH-Sequenzen, die davon kodierten Methionin-Synthasen sowie die funktionalen Homologen dieser Polynukleotide bzw. Proteine.

Gegenstand der Erfindung sind insbesondere auch die zur Durchführung obiger Verfahren notwendigen Expressionskonstrukte und Mikroorganismen.

Weitere Gegenstände der Erfindung sind somit insbesondere:

5

15

20

25

30

- das Plasmid pCIS lysC thr311ile, kodierend für lysC thr311ile oder ein funktionales Äquivalent davon, d.h. eine lysC-Mutante mit vergleichbarer, gegenüber dem Wildtyp erhöhter Aspartatkinase-Aktivität;
- ein Wirtsorganismus transformiert mit dem Plasmid pCIS lysC thr311ile, insbesondere ausgewählt unter Mikroorganismen der Gattung Corynebacterium, insbesondere der Art C. glutamicum, wie der transformierte Stamm LU1479 lysC 311ile;
- das Plasmid pC Phsdh metH_Sc, kodierend für metH aus Streptomyces coelicolor,
- ein Wirtsorganismus gemäß obiger Definition, transformiert mit einem Plasmid, kodierend für exgonenes metH; insbesondere transformiert mit dem Plasmid pC Phsdh metH_Sc;
- ein Wirtsorganismus gemäß obiger Definiiton mit Resistenz gegen wenigstens einen Methlonin-Biosynthesehemmstoff, wie der transformierte Stamm LU1479 lysC 311ile ET-16, gege-

8

benenfalls transformiert mit einer exogenen kodierenden metH Sequenz, wie der transformierte Stamm LU1479 lysC 311ile ET-16 pC Phsdh metH_Sc.

Detaillierte Beschreibung der Erfindung

5

10

15

20

a) Allgemeine Begriffe

Als Proteine mit der biologischen Aktivität der Methionine-Synthase, kurz auch als metH genannt (systematische Bezeichnung: 5-Methyltetrahydrofolat-Homocystein S-Methyltransferase :EC 2.1.1.13), werden solche Proteine bezeichnet, die in der Lage sind Homocystein unter Verwendung der Cofaktoren 5-Methyltetrahydrofolat (MTHF), Cobalamin (Vitamin B12) und S-Adenosyl-Methionin zu Methionin und Tetrahydrofolat umzusetzen. Während der Cofaktor 5-Methyltetrahydrofolat stöchiometrisch in die Reaktion mit eingeht (1mol MTHF/1Mol Methionin gebildet) wird, wie in der Literatur beschrieben, S-Adenosyl-Methionin substöchiometrisch umgesetzt. Cobalamin hingegen ist katalytisch an der Umsetzung beteiligt. Dem Fachmann sind weitere Details des metH-Proteins bekannt. (Banerjee R.V., Matthews R.G. FASEB J. 4:1450-1459, 1990, Ludwig ML. Matthews RG. Annual Review of Biochemistry. 66:269-313, 1997. Drennan CL. Matthews RG. Ludwig ML. Current Opinion in Structural Biology. 4:919-29, 1994). Der Fachmann unterscheidet die Aktivität der Cobalamin-abhängigen 5-Methyltetrahydrofolat-Homocystein S-Methyltransferase von der der Cobalamin-unabhängigen 5-Methyltetrahydro-Pteroyltriglutamat-Homocystein S-Methyltransferase (EC 2.1.1.14) auch metE genannt. Der Fachmann kann die enzymatische Aktivität von metH durch Enzymtests nachweisen, Vorschriften dafür können sein: Jarrett JT. Goulding CW. Fluhr K. Huang S. Matthews RG. Methods in Enzymology. 281:196-213, 1997.

25

30

Im Rahmen der vorliegenden Erfindung umfasst der Begriff "schwefelhaltige Feinchemikalie" jegliche chemische Verbindung, die wenigstens ein Schwefelatom kovalent gebunden enthält und durch ein erfindungsgemäßes Fermentationsverfahrens zugänglich ist. Nichtlimitierende Beispiele dafür sind Methionin, Homocystein, S-Adenosyl-Methionin, insbesondere Methionin, und S-Adenosyl-Methionin.

Im Rahmen der vorliegenden Erfindung umfassen die Begriffe L-Methionin, Methionin, Homocystein und S-Adenosylmethionin auch die korrespondierenden Salze, wie z. B. Methionin-Hydrochlorid oder Methionin-Sulfat.

5 "Polynukleotide" bezeichnet im allgemeinen Polyribonukleotide (RNA) und Polydeoxyribonukleotide (DNA), wobei es sich um nicht modifizierte RNA oder DNA oder modifizierte RNA oder DNA handeln kann.

Unter "Polypeptiden" versteht man erfindungsgemäß Peptide oder Proteine, die zwei oder mehr über Peptidbindungen verbundene Aminosäuren enthalten.

Der Begriff "Stoffwechselmetabolit" bezeichnet chemische Verbindungen, die im Stoffwechsel von Organismen als Zwischen- oder auch Endprodukte vorkommen und die neben ihrer Eigenschaft als chemische Bausteine auch modulierende Wirkung auf Enzyme und ihre katalytische 15 Aktivität haben können. Dabei ist aus der Literatur bekannt, dass solche Stoffwechselmetabolite sowohl hemmend als auch stimulierend auf die Aktvität von Enzymen wirken können (Biochemistry, Stryer, Lubert, 1995 W. H. Freeman & Company, New York, New York.). In der Literatur ist auch beschrieben, dass es möglich ist durch Maßnahmen wie Mutation der genomischen DNA durch UV-Strahlung, ionisierender Strahlung oder mutagene Substanzen und nachfolgen-20 der Selektion auf bestimmte Phänotypen in Organismen solche Enzyme zu produzieren, in denen die Beeinflussung durch Stoffwechselmetabolite verändert wurde (Sahm H. Eggeling L. de Graaf AA. Biological Chemistry 381(9-10):899-910, 2000; Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94). Diese veränderten Eigenschaften können auch durch gezielte Maßnahmen erreicht werden. Dabei ist dem Fachmann bekannt, in Genen für Enzyme auch gezielt bestimmte Nukleotide der für das Protein kodierenden DNA so zu verän-25 dem, dass das aus der exprimierten DNA-Sequenz resultierende Protein bestimmte neue Eigenschaften aufweist, so zum Beispiel, dass die modulierende Wirkung von Stoffwechselmetaboliten gegenüber dem nicht veränderten Protein verändert ist

30 Enzyme k\u00f6nnen derart in ihrer Aktivit\u00e4t beeinflu\u00dft werden, dass es zu einer Verringerung der Reaktionsgeschwindigkeit, oder zu einer Ver\u00e4nderung der Affinit\u00e4t gegen\u00fcber dem Substrat oder zu einer \u00e4nderung der Reaktionsgeschwindigkeiten kommt Die Begriffe "exprimieren" bzw. "Verstärkung" oder "Überexpression" beschreiben im Kontext der Erfindung die Produktion bzw. Erhöhung der intrazellulären Aktivität eines oder mehrerer Enzyme in einem Mikroorganismus, die durch die entsprechende DNA kodiert werden. Dazu kann man beispielsweise ein Gen in einen Organismus einbringen, ein vorhandenes Gen durch ein anderes Gen ersetzen, die Kopienzahl des Gens bzw. der Gene erhöhen, einen starken Promotor verwenden oder ein Gen verwenden, das für ein entsprechendes Enzym mit einer hohen Aktivität kodiert und man kann gegebenenfalls diese Maßnahmen kombinieren.

b) Erfindungsgemäße metH-Proteine

10

20

25

Erfindungsgemäß mit umfasst sind ebenfalls "funktionale Äquivalente" der konkret offenbarten metH-Enzyme aus Organismen obiger Liste I.

"Funktionale Äquivalente" oder Analoga der konkret offenbarten Polypeptide sind im Rahmen der vorliegenden Erfindung davon verschiedene Polypeptide, welche weiterhin die gewünschte biologische Aktivität, wie z.B. Substratspezifität, besitzen.

Unter "funktionalen Äquivalenten" versteht man erfindungsgemäß insbesondere Mutanten, welche in wenigstens einer der oben genannten Sequenzpositionen eine andere als die konkret genannte Aminosäure aufweisen aber trotzdem eine der oben genannten biologischen Aktivitäten besitzen. "Funktionale Äquivalente" umfassen somit die durch eine oder mehrere Aminosäure-Additionen, -Substitutionen, -Deletionen und/oder -Inversionen erhältlichen Mutanten, wobei die genannten Veränderungen in jeglicher Sequenzposition auftreten können, solange sie zu einer Mutante mit dem erfindungsgemäßen Eigenschaftsprofil führen. Funktionale Äquivalenz ist insbesondere auch dann gegeben, wenn die Reaktivitätsmuster zwischen Mutante und unverändertem Polypeptid qualitativ übereinstimmen, d.h. beispielsweise gleiche Substrate mit unterschiedlicher Geschwindigkeit umgesetzt werden.

"Funktionale Äquivalente" umfassen natürlich auch Polypeptide welche aus anderen Organis-30 men zugänglich sind, sowie natürlich vorkommende Varianten. Beispielsweise lassen sich durch Sequenzvergleich Bereiche homologer Sequenzregionen festlegen und in Anlehnung an die konkreten Vorgaben der Erfindung äquivalente Enzyme ermitteln.

11

"Funktionale Äquivalente" umfassen ebenfalls Fragmente, vorzugsweise einzelne Domänen oder Sequenzmotive, der erfindungsgemäßen Polypeptide, welche z.B. die gewünschte biologische Funktion aufweisen.

"Funktionale Äquivalente" sind außerdem Fusionsproteine, welche ein der oben genannten Polypeptidsequenzen oder davon abgeleitete funktionale Äquivalente und wenigstens eine weitere, davon funktionell verschiedene, heterologe Sequenz in funktioneller N- oder C-terminaler Verknüpfung (d.h. ohne gegenseitigen wesentliche funktionelle Beeinträchtigung der Fusionsproteinteile) aufweisen. Nichtlimitiernde Beispiele für derartige heterologe Sequenzen sind z.B. Signalpeptide, Enzyme, Immunoglobuline, Oberflächenantigene, Rezeptoren oder Rezeptorliganden.

Erfindungsgemäß mit umfasste "funktionale Äquivalente" sind Homologe zu den konkret offenbarten Proteinen. Diese besitzen, beispielsweise über die gesamte Länge, wenigstens 30%, oder etwa 40%, 50 %, vorzugsweise wenigstens etwa 60 %, 65%, 70%, oder 75% ins besondere wenigsten 85 %, wie z.B. 90%, 95% oder 99%, Homologie zu einer der konkret offenbarten Sequenzen, berechnet nach dem Algorithmus von Pearson und Lipman, Proc. Natl. Acad, Sci. (USA) 85(8), 1988, 2444-2448. Der Homologiegrad spiegelt insbesondere den Grad der Identität zwischen veränderter und nicht veränderter Sequenz wider.

20

15

Homologe der erfindungsgemäßen Proteine oder Polypeptide können durch Mutagenese erzeugt werden, z.B. durch Punktmutation oder Verkürzung des Proteins. Der Begriff "Homolog", wie er hier verwendet wird, betrifft auch eine variante Form des Proteins, die als Agonist oder Antagonist der Protein-Aktivität wirkt.

25

30

Homologe des erfindungsgemäßen Proteine können durch Screening kombinatorischer Banken von Mutanten, wie z.B. Verkürzungsmutanten, identifiziert werden. Beispielsweise kann eine variegierte Bank von Protein-Varianten durch kombinatorische Mutagenese auf Nukleinsäure-ebene erzeugt werden, wie z.B. durch enzymatisches Ligieren eines Gemisches synthetischer Oligonukleotide. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller Homologer aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheseautomaten durchgeführt werden, und das synthetische Gen kann dann in einen geeigneten Expressions-

WO 03/087386

vektor ligiert werden. Die Verwendung eines degenerierten Gensatzes ermöglicht die Bereitstellung sämtlicher Sequenzen in einem Gemisch, die den gewünschten Satz an potentiellen Proteinsequenzen codieren. Verfahren zur Synthese degenerierter Oligonukleotide sind dem Fachmann bekannt (Z.B. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).

Zusätzlich können Banken von Fragmenten des Protein-Codons verwendet werden, um eine variegierte Population von Protein-Fragmenten zum Screening und zur anschließenden Selektion von Homologen eines erfindungsgemäßen Proteins zu erzeugen. Bei einer Ausführungsform kann eine Bank von kodierenden Sequenzfragmenten durch Behandeln eines doppelsträngigen PCR-Fragmentes einer kodierenden Sequenz mit einer Nuklease unter Bedingungen, unter denen ein Nicking nur etwa einmal pro Molekül erfolgt, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, die Sense-/Antisense-Paare von verschiedenen genickten Produkten umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuclease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Durch dieses Verfahren kann eine Expressionsbank hergeleitet werden, die N-terminale, C-terminale und interne Fragmente mit verschiedenen Größen des erfidungsgemäßen Proteins kodiert.

20

25

30

10

15

Im Stand der Technik sind mehrere Techniken zum Screening von Genprodukten kombinatorischer Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und zum Screening von cDNA-Banken auf Genprodukte mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Screening der Genbanken anpassen, die durch kombinatorische Mutagenese erfindungsgemäßer Homologer erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Screening großer Genbanken, die einer Analyse mit hohem Durchsatz unterliegen, umfassen das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren der geeigneten Zellen mit der resultierenden Vektorenbank und Exprimieren der kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen codiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine Technik, die die Häufigkeit funktioneller Mutanten in den Banken vergrößert, kann in Kombination mit den Screeningtests verwendet

werden, um Homologe zu identifizieren (Arkin und Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331

c) Erfindungsgemäße Polynukleotide

5

Gegenstand der Erfindung sind ebenso Nukleinsäuresequenzen (einzel- und doppelsträngige DNA- und RNA-Sequenzen, wie z.B. cDNA und mRNA), kodierend für eines der obigen metH-Enzyme und deren funktionalen Äquivalenten, welche z.B. auch unter Verwendung künstlicher Nukleotidanaloga zugänglich sind.

10

Die Erfindung betrifft sowohl isolierte Nukleinsäuremoleküle, welche für erfindungsgemäße Polypeptide bzw. Proteine oder biologisch aktive Abschnitte davon kodieren, sowie Nukleinsäurefragmente, die z.B. zur Verwendung als Hybridisierungssonden oder Primer zur Identifizierung oder Amplifizierung von erfindungsgemäßer kodierenden Nukleinsäuren verwendet werden können.

15 nen.

Die erfindungsgemäßen Nukleinsäuremoleküle können zudem untranslatierte Sequenzen vom 3'- und/oder 5'-Ende des kodierenden Genbereichs enthalten

.

20

Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure zugegen sind und kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert wird.

25

Die Erfindung umfasst weiterhin die zu den konkret beschriebenen Nukleotidsequenzen komplementären Nukleinsäuremoleküle oder einen Abschnitt davon.

30

Die erfindungsgemäß Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von homologer Sequenzen in anderen Zelltypen und Organismen verwendbar sind. Solche Sonden bzw. Primer umfassen gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise mindestens etwa 25, wie z.B. etwa 40, 50 oder 75 aufeinanderfolgende Nukleotide eines

15

30

Sense-Stranges einer erfindungsgemäßen Nukleinsäuresequenz oder eines entsprechenden Antisense-Stranges hybridisiert.

Weitere erfindungsgemäße Nukleinsäuresequenzen sind abgeleitet von SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49 oder 51 und unterscheiden sich davon durch Addition, Substitution, Insertion oder Deletion einzelner oder mehrerer Nukleotide, kodieren aber weiterhin für Polypeptide mit dem gewünschten Eigenschaftsprofil. Dies können Polynukleotide sein, die zu obigen Sequenzen, beispielsweise über die gesamte Länge, in mindestens etwa 50%, 55%, 60%, 65%, 70%, 80% oder 90%, vorzugsweise in mindestens etwa 95%, 96%, 97%, 98% oder 99% der Sequenzpositionen identisch sind.

Erfindungsgemäß umfasst sind auch solche Nukleinsäuresequenzen, die sogenannte stumme Mutationen umfassen oder entsprechend der Codon-Nutzung eins speziellen Ursprungs- oder Wirtsorganismus, im Vergleich zu einer konkret genannten Sequenz verändert sind, ebenso wie natürlich vorkommende Varianten, wie z.B. Spleißvarianten oder Allelvarianten, davon. Gegenstand sind ebenso durch konservative Nukleotidsubstutionen (d.h. die betreffende Aminosäure wird durch eine Aminosäure gleicher Ladung, Größe, Polarität und/oder Löslichkeit ersetzt) erhältliche Sequenzen.

20 Gegenstand der Erfindung sind auch die durch Sequenzpolymorphismen von den konkret offenbarten Nukleinsäuren abgeleiteten Moleküle. Diese genetischen Polymorphismen können zwischen Individuen innerhalb einer Population aufgrund der natürlichen Variation existieren. Diese natürlichen Variationen bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz eines Gens.

Weiterhin umfasst die Erfindung auch Nukleinsäuresequenzen, welchen mit oben genannten kodierenden Sequenzen hybridisieren oder dazu komplementär sind. Diese Polynukleotide lassen sich bei Durchmusterung von genomischen oder cDNA-Banken auffinden und gegebenenfalls daraus mit geeigneten Primern mittels PCR vermehren und anschließend beispielsweise mit geeigneten Sonden isolieren. Eine weitere Möglichkeit bietet die Transformation geeigneter Mikroorganismen mit erfindungsgemäßen Polynukleotiden oder Vektoren, die Vermehrung der Mikroorganismen und damit der Polynukleotide und deren anschließende Isolierung. Darüber hinaus können erfindungsgemäße Polynukleotide auch auf chemischem Wege synthetisiert

werden.

Unter der Eigenschaft, an Polynukleotide "hybridisieren" zu können, versteht man die Fähigkeit eines Poly- oder Oligonukleotids unter stringenten Bedingungen an eine nahezu komplementäre 5 Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen zu 70-100%, vorzugsweise zu 90-100%, komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northem- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze. Üblicherweise werden dazu Oligonukleotide ab einer Länge von 30 Basenpaaren eingesetzt. Unter stringenten Bedingungen versteht man beispielsweise in der Northern-Blot-Technik die Verwendung einer $50-70\,^{\circ}\text{C}$, vorzugsweise $60-65\,^{\circ}\text{C}$ warmen Waschlösung, beispielsweise $0.1x\,\text{SSC-Puffer mit}$ -0,1% SDS (20x SSC: 3M NaCl, 0,3M Na-Citrat, pH 7,0) zur Elution unspezifisch hybridisierter cDNA-Sonden oder Oligonukleotide. Dabei bleiben, wie oben erwähnt, nur in hohem Maße komplementäre Nukleinsäuren aneinander gebunden. Die Einstellung stringenter Bedingungen ist dem Fachmann bekannt und ist z:B. in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

c) Isolierung der kodierenden metH-Gene

20

15

Die für das Enzym Methionin-Synthase(EC 2.1.1.13) kodierenden metH Gene aus den Organismen obiger Liste I sind in an sich bekannter Weise isolierbar.

Zur Isolierung der metH-Gene oder auch anderer Gene der Organismen aus obiger Liste I wird zunächst eine Genbank dieses Organsimus in Escherichia coli (E. coli) angelegt. Das Anlegen 25 von Genbanken ist in allgemein bekannten Lehrbüchern und Handbüchern ausführlich beschrieben. Als Beispiel seien das Lehrbuch von Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Deutschland, 1990), oder das Handbuch von Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 30 1989) genannt. Eine sehr bekannte Genbank ist die des E. coli K-12 Stammes W3110, die von Kohara et al. (Cell50, 495-508 (198)) in λ-Vektoren angelegt wurde.

Zur Herstellung einer Genbank von Organismen der Liste I in E. coli können Cosmide, wie der Cosmidvektor SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84: 2160-2164), aber auch Plasmide, wie pBR322 (BoliVal; Life Sciences, 25, 807-818 (1979)) oder pUC9 (Vieira et al., 1982, Gene, 19: 259-268), verwendet werden. Als Wirte eignen sich besonders solche E. coli Stämme, die restriktions- und rekombinationsdefekt sind. Ein Beispiel hierfür ist der Stamm DH5αmcr, der von Grant et al. (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649) beschrieben wurde. Die mit Hilfe von Cosmiden klonierten langen DNA-Fragmente können anschließend wiederum in gängige, für die Sequenzierung geeignete Vektoren subkloniert und anschließend sequenziert werden, so wie es z. B. bei Sanger et al. (proceedings of the National Academy of Sciences of the United States of America, 74: 5463-5467, 1977) beschrieben ist.

Die erhaltenen DNA-Sequenzen können dann mit bekannten Algorithmen bzw. Sequenzanalyse-Programmen, wie z. B. dem von Staden (Nucleic Acids Research 14,217-232(1986)), dem von Marck (Nucleic Acids Research 16, 1829-1836 (1988)) oder dem GCG-Programm von Butler (Methods ofBiochemical Analysis 39, 74-97 (1998)), untersucht werden.

Die für die metH-Gene kodierenden DNA-Sequenzen von Organismen gemäß obiger Liste I wurden gefunden. Insbesondere wurden DNA-Sequenzen gemäß gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49 und 51 gefunden. Weiterhin wurde aus diesen vorliegenden DNA-Sequenzen mit den oben beschriebenen Methoden die Aminosäuresequenzen der entsprechenden Proteine abgeleitet. Durch SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 und 52 sind die sich ergebenden Aminosäuresequenzen der metH Genprodukte dargestellt.

25

30

10

15

20

Kodierende DNA-Sequenzen, die sich aus den Sequenzen gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49 und 51 durch die Degeneration des genetischen Kodes ergeben, sind ebenfalls Gegenstand der Erfindung. In gleicher Weise sind DNA-Sequenzen, die mit diesen Sequenzen oder davon abgeleiteten Sequenzteilen hybridisieren, Gegenstand der Erfindung.

Anleitungen zur Identifizierung von DNA-Sequenzen mittels Hybridisierung findet der Fachmann unter anderem im Handbuch "The DIG System Users Guide für Filter Hybridization" der Firma

Boehringer Mannheim GmbH (Mannheim, Deutschland, 1993) und bei Liebl et al. (International Journal of Systematic Bacteriology (1991) 41: 255-260). Anleitungen zur Amplifikation von DNA-Sequenzen mit Hilfe der Polymerase-Kettenreaktion (PCR) findet der Fachmann unter anderem im Handbuch von Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press, Ox-ford, UK, 1984) und bei Newton und Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Deutschland, 1994).

Weiterhin ist bekannt, dass Änderungen am N- und/oder C- Terminus eines Proteins dessen Funktion nicht wesentlich beeinträchtigen oder sogar stabilisieren können. Angaben hierzu findet der Fachmann unter anderem bei Ben-Bassat et al. (Journal of Bacteriology 169: 751-757 (1987)), bei O'Regan et al. (Gene 77: 237-251 (1989), bei Sahin-Toth et al. (Protein Sciences 3: 240–247 (1994)), bei Hochuli et al. (Biontechnology 6: 1321-1325 (1988)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.

Aminosäuresequenzen, die sich in entsprechender Weise aus den SEQ ID NO:2, 4, 6, 8, 10, 12,
 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 und 52 ergeben, sind ebenfalls Bestandteil der Erfindung.

d) <u>Erfindungsgemäß verwendete Wirtszellen</u>

20

10

Weitere Gegenstände der Erfindung betreffen als Wirtszelle dienende Mikroorgansismen, insbesondere coryneforme Bakterien, die einen Vektor, insbesondere Pendelvektor oder Plasmidvektor, der wenigstens ein metH Gen gerfindungsgemäßer Definition trägt, enthalten oder in denen ein erfindungsgemäßes metH Gen exprimiert bzw. verstärkt ist.

25

Diese Mikroorganismen können schwefelhaltige Feinchemikalien, insbesondere L-Methionin, aus Glucose, Saccharose, Lactose, Fructose, Maltose, Melasse, Stärke, Cellulose oder aus Glycerin und Ethanol herstellen. Vorzugsweise sind dies coryneforme Bakterien, insbesondere der Gattung Corynebacterium. Aus der Gattung Corynebacterium ist insbesondere die Art Corynebacterium glutamicum zu nennen, die in der Fachwelt für ihre Fähigkeit bekannt ist, L-Aminosäuren zu produzieren.

Als Beispiele für geeignete Stämme coryneformer Bakterien sind solche der Gattung Corynebacterium, insbesondere der Art Corynebacterium glutamicum (C. glutamicum), wie Corynebacterium glutamicum ATCC 13032,

Corynebacterium acetoglutamicum ATCC 15806,

5 Corynebacterium acetoacidophilum ATCC 13870, Corynebacterium thermoaminogenes FERM BP-1539, Corynebacterium melassecola ATCC 17965

oder

der Gattung Brevibacterium, wie
 Brevibacterium flavum ATCC 14067
 Brevibacterium lactofermentum ATCC 13869 und
 Brevibacterium divaricatum ATCC 14020 zu nennen;
 oder davon abgeleitete Stämme, wie
 Corynebacterium glutamicum KFCC10065

Corynebacterium glutamicum ATCC21608

welche ebenfalls die gewünschte Feinchemikalie oder deren Vorstufe(n) produzieren. Mit der Abkürzung KFCC ist die Korean Federation of Culture Collection gemeint, mit der Abkürzung ATCC die American type strain culture collection und mit der Abkürzung FERM die Sammlung des National institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Japan.

e) <u>Durchführung der erfindungsgemäßen Fermentation</u>

25

20

Erfindungsgemäß wurde festgestellt, dass coryneforme Bakterien nach Überexpression eines metH Gens aus Organismen der Liste I in vorteilhafter Weise schwefelhaltige Feinchemikalien, insbesondere L-Methionin, produzieren.

30 Zur Erzielung einer Überexpression kann der Fachmann unterschiedliche Maßnahmen einzeln oder in Kombination ergreifen. So kann die Kopienzahl der entsprechenden Gene erhöht werden, oder es kann die Promotor- und Regulationsregion oder die Ribosomenbindungsstelle, die sich stromaufwärts des Strukturgens befindet, mutiert werden. In gleicher Weise wirken Expres-

19

sionskassetten, die stromaufwärts des Strukturgens eingebaut werden. Durch induzierbare Promotoren ist es zusätzlich möglich, die Expression im Verlaufe der fermentativen L-Methionin-Produktion zu steigem. Durch Maßnahmen zur Verlängerung der Lebensdauer der mRNA wird ebenfalls die Expression verbessert. Weiterhin wird durch Verhinderung des Abbaus des Enzymproteins ebenfalls die Enzymaktivität verstärkt. Die Gene oder Genkonstrukte können entweder in Plasmiden mit unterschiedlicher Kopienzahl vorliegen oder im Chromosom integriert und amplifiziert sein. Alternativ kann weiterhin eine Überexpression der betreffenden Gene durch Veränderung der Medienzusammensetzung und Kulturführung erreicht werden.

Anleitungen hierzu findet der Fachmann unter anderem bei Martin et al. (Biontechnology 5, 137-146 (1987)), bei Guerrero et al. (Gene 138, 35-41 (1994)), Tsuchiya und Morinaga (Bio/Technology 6, 428-430 (1988)), bei Eikmanns et al. (Gene 102, 93-98 (1991)), in der Europäischen Patentschrift 0472869, im US Patent 4,601,893, bei Schwarzer und Pühler (Biotechnology 9, 84-87 (1991), bei Remscheid et al. (Applied and Environmental Microbiology 60,126-132 (1994), bei LaBarre et al. (Journal of Bacteriology 175, 1001-1007 (1993)), in der Patentanmeldung WO 96/15246, bei Malumbres et al. (Gene 134, 15-24 (1993)), in der japanischen Offenlegungsschrift JP-A-10-229891, bei Jensen und Hammer (Biotechnology and Bioengineering 58,.191-195 (1998)), bei Makrides (Microbiological Reviews 60 : 512-538 (1996) und in bekannten Lehrbüchem der Genetik und Molekularbiologie.

20

25

30

Gegenstand der Erfindung sind deshalb auch Expressionskonstrukte, enthaltend unter der genetischen Kontrolle regulativer Nukleinsäuresequenzen eine für ein erfindungsgemäßes Polypeptid kodierende Nukleinsäuresequenz; sowie Vektoren, umfassend wenigstens eines dieser Expressionskonstrukte. Vorzugsweise umfassen solche erfindungsgemäßen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils operativ verknüpft mit der kodierenden Sequenz. Unter einer "operativen Verknüpfung" versteht man die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und gegebenenfalls weiterer regulativer Elemente derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann. Beispiele für operativ verknüpfbare Sequenzen sind Aktivrieungssequenzen sowie Enhancer und dergleichen. Weitere regulative Elemente umfassen selektierbare Marker, Amplifikationssignale, Replikationsursprünge und dergleichen. Geeignete regulatorische Sequenzen sind z.B. be-

schrieben in Goeddel, Gene Expression Technology. Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

Zusätzlich zu den artifiziellen Regulationssequenzen kann die natürliche Regulationssequenz vor dem eigentlichen Strukturgen noch vorhanden sein. Durch genetische Veränderung kann diese natürliche Regulation gegebenenfalls ausgeschaltet und die Expression der Gene erhöht oder erniedrigt werden. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es werden keine zusätzlichen Regulationssignale vor das Strukturgen insertiert und der natürliche Promotor mit seiner Regulation wird nicht entfemt. Statt dessen wird die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert oder verringert wird. Die Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

Beispiele für brauchbare Promotoren sind: die Promotoren, ddh, amy, lysC, dapA, lysA aus Corynebacterium glutamicum, aber auch gram-positiven Promotoren SPO2 wie sie in Bacillus Subtilis and Its Closest Relatives, Sonenshein, Abraham L.,Hoch, James A., Losick, Richard; ASM Press, District of Columbia, Washington und Patek M. Eikmanns BJ. Patek J. Sahm H. Microbiology. 142 1297-309, 1996 beschrieben sind, oder aber auch cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, laclq-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-, lambda-PR- oder lambda-PL-Promotor, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Bevorzugt ist auch die Verwendung induzierbarer Promotoren, wie z.B. licht- und insbesondere temperaturinduztierbarer Promotoren, wie der PrP-Promotor. Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

25

10

15

20

Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen ermöglichen. Dies kann beispielswelse je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird.

30

Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke

25

30

Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors, einer geeigneten Shine-Dalgamo-Sequenz mit einer metH-Nukleotidsequenz sowie einem geeigneten Terminationssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in Current Protocols in Molecular Biology, 1993, John Wiley & Sons, Incorporated, New York New York, PCR Methods, Gelfand, David H., Innis, Michael A., Sninsky, John J. 1999, Academic Press, Incorporated, California, San Diego, ., PCR Cloning Protocols, Methods in Molecular Biology Ser., Vol. 192, 2nd ed., Humana Press, New Jersey, Totowa. T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory,
 Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.

Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor insertiert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannten Vektoren, wie beispielsweise Phagen, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.

Zur Verstärkung wurden erfindungsgemäße metH Gene beispielhaft mit Hilfe von episomalen Plasmiden überexprimiert. Als Plasmide eignen sich solche, die in coryneformen Bakterien repliziert werden. Zahlreiche bekannte Plasmidvektoren, wie z. B. pZ1 (Menkel et al., Applied and Environmental Microbiology (1989) 64: 549-554), pEKEx1 (Eikmanns et al., Gene 102: 93-98 (1991)) oder pHS2-1 (Sonnen et al., Gene 107: 69-74 (1991)) beruhen auf den kryptischen Plasmiden pHM1519, pBL1 oder pGA1. Andere Plasmidvektoren, wie z. B. pCLiK5MCS, oder solche, die auf pCG4 (US-A 4,489,160) oder pNG2 (Serwold-Davis et al., FEMS Microbiology

Letters 66, 119-124 (1990)) oder pAG1 (US-A 5,158,891) beruhen, können in gleicher Weise verwendet werden.

Weiterhin eignen sich auch solche Plasmidvektoren mit Hilfe derer man das Verfahren der Genamplifikation durch Integration in das Chromosom anwenden kann, so wie es belspielsweise von Remscheid et al. (Applied and Environmental Microbiology 60,126-132 (1994)) zur Duplikation bzw. Amplifikation des hom-thrB-Operons beschrieben wurde. Bei dieser Methode wird das vollständige Gen in einen Plasmidvektor kloniert, der in einem Wirt (typischerweise E. coli), nicht aber in C. glutamicum replizieren kann. Als Vektoren kommen beispielsweise pSUP301 (Simon et al., Bio/ Technology 1,784-791 (1983)), pK18mob oder pK19mob (Schäfer et al., Gene 145,69-73 (1994)), Bernard et al., Journal ofMolecular Biology, 234: 534-541 (1993)), pEM1 (Schrumpf et al. 1991, Journal of Bacteriology 173: 4510–4516) oder pBGS8 (Spratt et al.,1986, Gene 41: 337-342) in Frage. Der Plasmidvektor, der das zu amplifizierende Gen enthält, wird anschließend durch Transformation in den gewünschten Stamm von C. glutamicum überführt. Methoden zur Transformation sind beispielsweise bei Thierbach et al. (Applied Microbiology and Biotechnology 29, 356-362 (1988)), Dunican und Shivnan (Biotechnology 7, 1067-1070 (1989)) und Tauch et al. (FEMS Microbiological Letters 123,343-347 (1994)) beschrieben.

Enzyme können durch Mutationen in den korrespondierenden Genen derart in ihrer Aktivität beeinflußt werden, dass es zu einer teilweisen oder vollständigen Verringerung der Reaktionsgeschwindigkeit der enzymatischen Reaktion kommt. Beispiele für solche Mutationen sind dem Fachmann bekannt (Motoyama H. Yano H. Terasaki Y. Anazawa H. Applied & Environmental Microbiology. 67:3064-70, 2001, Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94.)

25

30

10

15

20

Zusätzlich kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben einer Expression bzw. Verstärkung eines erfindungsgemäßen metH-Gen eines oder mehrere Enzyme des Methionin-Biosyntheseweges oder eines damit assozlierten (d.h. in einem funktionelle Zusammenhang stehenden) Biosynthese – oder sonstigen Stoffwechselweges, wie des Cystein-, Lysin- oder Threonin-Stoffwechselwegs, wie insbesondere der Aspratatsemialdehyd-Synthese, der Glykolyse, der Anaplerotik, des Pentose-Phosphat-Stoffwechsels, des Zitronensäure-Zyklus oder des Aminosäure-Exports zu verstärken.

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, eines oder mehrere der folgenden Gene verstärkt sein, (d.h. z.B. in einer höheren Koplenzahl vorliegen oder ein Enzym mit höherer Aktivität oder Spezifität kodieren):

- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),
- -das für eine Aspartat-Semialdehyd Dehydrogenasekodierende Gen asd (EP 1 108 790 A2;
 DNA-SEQ NO. 282),
 - das für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierende Gen gap (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die 3-Phosphoglycerat Kinase kodierende Gen pgk (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
 - das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
 - das für die Triosephosphat Isomerase kodierende Gen tpi (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das f
 ür die Homoserin O-Acetyltransferase kodierende Gen metA (EP 1 108 790 A2; DNA-SEQ
 NO. 725),
 - das für die Cystahionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),
 - das für die Cystahionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),
 - das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
 - das für die O-Acetylhomoserin-Sulfhydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),
- das f
 ür die Methylentetrahydrofolat-Reduktase kodierende Gen met
 EP 1 108 790 A2; DNA SEQ NO. 2379),
 - das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)
- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ
 NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)
 - das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO. 2818)

- das für eine Homosenn-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)
- So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, in coryneformen Bakterien, vorteilhaft sein, gleichzeitig wenigstens eines der nachfolgenden Gene zu mutieren, insbesondere so, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch einen Stoffwechselmetaboliten in ihrer Aktivität beeinflusst werden:
- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281).
 - das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
 - das für die Homoserin O-Acetyltransferase kodierende Gen metA (EP 1 108 790 A2; DNA-SEQ NO. 725),
- das für die Cystahionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ
 NO. 3491),
 - das für die Cystahionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),
- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
 - das für die O-Acetylhomoserin-Sulfhydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),
 - das für die Methylentetrahydrofolat-Reduktase kodierende Gen metF (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)
 - eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)
 - das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO.
- 30 2818)
 - das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen metH-Gene eines oder mehrere der folgenden Gene abzuschwächen, insbesondere deren Expression zu verringern, oder auszuschalten:

5

20

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)
- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)
- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2;
 DNA-SEQ NO. 3494)
 - das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)
- das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)
 - das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)
 - das für die Dihydrodipicolinat Synthase kodiemde Gen dapA(EP 1 108 790 A2; DNA-SEQ NO. 3476)
 - das für die Dihydrodipicolinat Reduktase kodiemde Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)
 - das für die Diaminopicolinat Decarboxylase kodiemde Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere LMethionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen metH-Gene in Coryneformen Bakterien gleichzeitig wenigstens eines der folgenden Gene so zu mutieren, dass die enzymatische Aktivität des korrespondierenden Proteins teilweise
oder vollständig verringert wird:

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)
 - das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)
 - das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)

- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)
- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)
- das f
 ür die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)
 - das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)
 - das für die Dihydrodipicolinat Synthase kodiernde Gen dapA(EP 1 108 790 A2; DNA-SEQ NO. 3476)
- das für die Dihydrodipicolinat Reduktase kodiernde Gen dapB (EP 1 108 790 A2; DNA-SEQ
 NO. 3477)
 - das für die Diaminopicolinat Decarboxylase kodiemde Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)
- Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben der Expression bzw. Verstärkung eines erfindungsgemäßen metH-Gens unerwünschte Nebenreaktionen auszuschalten, welche beispielsweise die Ausbeute an der Feichchemikalie verringem (Nakayama: "Breeding of Amino Acid Producing Microorganisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).
 - Die erfindungsgemäß hergestellten Mikroorganismen können kontinuierlich oder diskontinuierlich im batch- Verfahren (Satzkultivierung) oder im fed batch (Zulaufverfahren) oder repeated fed
 batch Verfahren (repetitives Zulaufverfahren) zur Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, kultiviert werden. Eine Zusammenfassung über bekannte Kultivierungsmethoden ist im Lehrbuch von Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) oder im Lehrbuch von Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)) zu finden.
- 30 Das zu verwendende Kulturmedium hat in geeigneter Welse den Ansprüchen der jeweiligen Stämme zu genügen. Beschreibungen von Kulturmedien verschiedener Mikroorganismen sind im Handbuch "Manual of Methods für General Bacteriology" der American Society für Bacteriology (Washington D. C., USA, 1981) enthalten.

Diese erfindungsgemäß einsetzbaren Medien umfassen gewöhnlich eine oder mehrere Kohlenstoffquellen, Stickstoffquellen, anorganische Salze, Vitamine und/oder Spurenelemente.

- Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind beispielsweise Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cellulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte der Zucker-Raffinierung zu den Medien geben. Es kann auch vorteilhaft sein, Gemische verschledener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Öle und Fette wie z. B. Sojaöl. Sonnenblumenöl. Erdnußöl und Kokosfett, Fettsäuren wie z. B. Palmitinsäure, Stearinsäure oder Linolsäure, Alkohole wie z. B. Glycerin, Methanol oder Ethanol und organische Säuren wie z. B. Essigsäure oder Milchsäure.
- Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffverbindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stickstoffquellen umfassen Ammoniak-Gas oder Ammoniumsalze, wie Ammoniumsulfat, Ammoniumchlorid, Ammoniumphosphat, Ammoniumcarbonat oder Ammoniumnitrat, Nitrate, Hamstoff, Aminosäuren oder komplexe Stickstoffquellen, wie Maisquellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und andere. Die Stickstoffquellen können einzeln oder als Mischung verwendet werden.

Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen die Chlorid-, Phosphor- oder Sulfatsalze von Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen

25

Als Schwefelquelle für die Herstellung von schwefelhaltigen Feinchemikalien, insbesondere von Methionin, können anorganische schwefelhaltige Verbindungen wie beispielsweise Sulfate, Sulfite, Dithionite, Tetrathionate, Thiosulfate, Sulfide aber auch organische Schwefelverbindungen, wie Mercaptane und Thiole, verwendet werden.

30

Als Phosphorquelle können Phosphorsäure, Kaliumdihydrogenphosphat oder Dikaliumhydrogenphosphat oder die entsprechenden Natrium haltigen Salze verwendet werden.

28

Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu halten. Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol oder Protocatechuat, oder organische Säuren, wie Citronensäure.

Die erfindungsgemäß eingesetzten Fermentationsmedien enthalten üblicherweise auch andere Wachstumsfaktoren, wie Vitamine oder Wachstumsförderer, zu denen beispielsweise Biotin, Riboffavin, Thiamin, Folsäure, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Melassen, Maisquellwasser und dergleichen. Dem Kulturmedium können überdies geeignete Vorstufen zugesetzt werden. Die genaue Zusammensetzung der Medienverbindungen hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individuell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehrbuch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wachstumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.

Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der Anzucht zugegen sein oder wahlfrei kontinuierlich oder chargenweise hinzugegeben werden.

20

25

30

Die Temperatur der Kultur liegt normalerweise zwischen 15°C und 45°C, vorzugsweise bei 25°C bis 40°C und kann während des Experimentes konstant gehalten oder verändert werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen. Der pH-Wert für die Anzucht läßt sich während der Anzucht durch Zugabe von basische Verbindungen wie Natriumhydroxid, Kaliumhydroxid, Ammoniak bzw. Ammoniakwasser oder saure Verbindungen wie Phosphorsäure oder Schwefelsäure kontrollieren. Zur Kontrolle der Schaumentwicklung können Antischaummitte,I wie z. B. Fettsäurepolyglykolester, eingesetzt werden. Zur Aufrechterhaltung der Stabilität von Plasmiden können dem Medium geeignete selektiv wirkende Stoffe, wie z. B. Antibiotika, hinzugefügt werden. Um aerobe Bedingungen aufrechtzuerhalten, werden Sauerstoff oder Sauerstoff haltige Gasmischungen, wie z. B. Umgebungsluft, in die Kultur eingetragen. Die Temperatur der Kultur liegt normalerweise bei 20°C bis 45°C. Die Kultur wird solange fortgesetzt, bis sich ein Maximum des gewünschten Produktes gebildet hat. Dieses Ziel wird

normalerweise innerhalb von 10 Stunden bis 160 Stunden erreicht.

Die so erhaltenen, insbesondere L-Methionin enthaltenden, Fermentationsbrühen haben üblicherweise eine Trockenmasse von 7,5 bis 25 Gew.-%.

5

Vorteilhaft ist außerdem auch, wenn die Fermentation zumindest am Ende, insbesondere jedoch über mindestens 30% der Fermentationsdauer zuckerlimitiert gefahren wird. Das heißt, dass während dieser Zeit die Konzentration an verwertbarem Zucker im Fermentationsmedium auf ≥ 0 bis 3 g/l gehalten, beziehungsweise abgesenkt wird.

10

Die Fermentationsbrühe wird anschließend weiterverarbeitet. Je nach Anforderung kann die Biomasse ganz oder teilweise durch Separationsmethoden, wie z. B. Zentrifugation, Filtration, Dekantieren oder einer Kombination dieser Methoden aus der Fermentationsbrühe entfemt oder vollständig in ihr belassen werden.

15

20

Anschließend kann die Fermentationsbrühe mit bekannten Methoden, wie z. B. mit Hilfe eines Rotationsverdampfers, Dünnschichtverdampfers, Fallfilmverdampfers, durch Umkehrosmose, oder durch Nanofiltration, eingedickt beziehungsweise aufkonzentriert werden. Diese aufkonzentrierte Fermentationsbrühe kann anschließend durch Gefriertrocknung, Sprühtrocknung, Sprühtrocknung, Sprühtranulation oder durch anderweitige Verfahren aufgearbeitet werden.

25

Es ist aber auch möglich die schwefelhaltigen Feinchemikalien, insbesonder L-Methionin, weiter aufzureinigen. Hierzu wird die produkthaltige Brühe nach dem Abtrennen der Biomasse einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Produkt oder die Verunreinigungen ganz oder teilweise auf dem Chromatographieharz zurückgehalten werden. Diese Chromatographieschnitte können nötigenfalls wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl der geeigneten Chromatographieharze und ihrer wirksamsten Anwendung bewandert. Das gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

30

Die Identität und Reinheit der isolierten Verbindung(en) kann durch Techniken des Standes der Technik bestimmt werden. Diese umfassen Hochleistungs-Flüssigkeitschromatographie (HPLC),

spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, NIRS, Enzymtest oder mikrobiologische Tests. Diese Analyseverfahren sind zusammengefaßt in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotekhnologiya 11 27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Bd. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

10 Die Erfindung wird nun anhand der folgenden nicht-limitierenden Beispiele näher beschrieben:

Beispiel 1: Konstruktion von pCLiK5MCS

Zunächst wurden Ampicillinresistenz und Replikationsursprung des Vektors pBR322 mit den Oligonukleotiden p1.3 (SEQ ID NO:53) und p2.3 (SEQ ID NO:54) mit Hilfe der Polymerase-Kettenreaktion (PCR) amplifiziert.

p1.3 (SEQ ID NO:53)

5'-CCCGGGATCCGCTAGCGGCGGCCGGCCGGCCGGTGTGAAATACCGCACAG-3'

20

25

30

p2.3 (SEQ ID NO:54)

5'-TCTAGACTCGAGCGGCCGGCCGGCCTTTAAATTGAAGACGAAAGGGCCTCG-3'

Neben den zu pBR322 komplementären Sequenzen, enthält das Oligonukleotid p1.3 (SEQ ID NO:53) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen Smal, BamHI, Nhel und Ascl und das Oligonukleotid p2.3 (SEQ ID NO:54) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen Xbal, XhoI, NotI und Dral. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,1 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Die stumpfen Enden des DNA-Fragmentes wurden mit dem Rapid DNA Ligation Kit (Roche

Diagnostics, Mannheim) nach Angaben des Herstellers miteinander ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK1.

10

Ausgehend vom Plasmid pWLT1 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden neo1 (SEQ ID NO:55) und neo2 (SEQ ID NO:56) eine Kanamycin-Resistenzcassette amplifiziert.

15 neo1 (SEQ ID NO:55):

5'-GAGATCTAGACCCGGGGATCCGCTAGCGGGCTGCTAAAGGAAGCGGA-3'

neo2 (SEQ ID NO:56):

5'-GAGAGGCGCCGCTAGCGTGGGCGAAGAACTCCAGCA-3'

20

25

30

Neben den zu pWLT1 komplementären Sequenzen, enthält das Oligonukleotid neo1 in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen Xbal, Smal, BamHl, Nhel und das Oligonukleotid neo2 (SEQ ID NO:56) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen Ascl und Nhel. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,3 kb wurde mit dem GFXTPCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit den Restriktionsendonukleasen Xbal und Ascl (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran emeut mit dem GFXTPCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK1 wurde ebenfalls mit den Restriktionsendonukleasen Xbal und Ascl geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim) nach Angaben des Herstellers

dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,1kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50μg/ml) und Kanamycin (20μg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

10

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK2.

Der Vektor pCLiK2 wurde mit der Restriktionsendonuklease Dral (New England Biolabs, Beverty, USA) geschnitten. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ca. 2,3 kb großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers religiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20μg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

25

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK3.

30 Ausgehend vom Plasmid pWLQ2 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden cg1 ((SEQ ID NO:57) und cg2 (SEQ ID NO:58) der Replikationsursprung pHM1519 amplifiziert.

cg1 (SEQ ID NO:57): 5'-GAGAGGCGGCCGCGCAAAGTCCCGCTTCGTGAA-3'

cg2 (SEQ ID NO:58):

30

5 5'-GAGAGGGCGGCCGCTCAAGTCGGTCAAGCCACGC-3'

Neben den zu pWLQ2 komplementären Sequenzen, enthalten die Oligonukleotide cg1 (SEQ ID NO:57) und cg2 (SEQ ID NO:58) Schnittstellen für die Restriktionsendonuklease Notl. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and 10 Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,7 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease Notl (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem 15 GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK3 wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese In einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,3kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit 20 (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment liglert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) 25 transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5.

Für die Erweiterung von pCLik5 um eine "multiple cloning site" (MCS) wurden die beide synthetischen, weitestgehend komplementären Oligonukleotide HS445 ((SEQ ID NO:59) und HS446

(SEQ ID NO:60), die Schnittstellen für die Restriktionsendonukleasen Swal, Xhol, Aatl, Apal, Asp718, Mlul, Ndel, Spel, EcoRV, Sall, Clal, BamHI, Xbal und Smal enthalten, durch gemeinsames erhitzen auf 95°C und langsames abkühlen zu einem doppelsträngigen DNA-Fragment vereinigt.

5

HS445 (SEQ ID NO:59):

10

HS446 (SEQ ID NO:60):

15

-20

25

Der Vektor pCLiK5 wurde mit den Restriktionsendonuklease Xhol und BamHI (New England Biolabs, Beverly, USA) geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8% Igen Agarosegel wurde der linearisierte Vektor (ca. 5,0 kb) mit dem GFXTMPCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem synthetischen Doppelsträngigen DNA-Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Aca-

35

demy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS ist als SEQ ID NO: 63 aufgeführt.

5

10

20

25

Beispiel 2: Konstruktion von pCLiK5MCS integrativ sacB

Ausgehend vom Plasmid pK19mob (Schäfer et al., Gene 145,69-73(1994)) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden BK1732 und BK1733 das Bacillus subtilis sacB Gen (kodierend für Levan Sucrase) amplifiziert.

BK1732 (SEQ ID NO:61): 5'-GAGAGCGGCCGCCGATCCTTTTTAACCCATCAC-3'

15 BK1733 (SEQ ID NO:62): 5'-AGGAGCGGCCGCCATCGGCATTTTCTTTTGCG-3'

Neben den zu pEK19mobsac komplementären Sequenzen, enthalten die Oligonukleotide BK1732 und BK1733 Schnittstellen für die Restriktionsendonuklease Notl. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,9 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease Notl (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

Der Vektor pCLiK5MCS (hergestellt gemäß Beispiel 1) wurde ebenfalls mit der Restriktionsendonuklease Notl geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen
Agarosegel wurde ein ungefähr 2,4 kb großes Vektorfragment mit dem GFX™PCR, DNA and
Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers iso-

liert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS integrativ sacB.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

15

10

Das entstandene Plasmid pCLiK5MCS integrativ sacB ist als SEQ ID NO: 64 aufgeführt.

Weitere Vektoren die zur erfindungsgemäßen Expression oder Überproduktion von metH-Genen geeignet sind, können in analoger Weise herstellt werden.

20

30

In den folgenden Beispielen 3 bis 8 wird die schrittweise Konstruktion eines verbesserten Methionin-produzierenden Stammes mit der Bezeichnung LU1479 lysC 311ile ET-16 pC Phsdh metH_Sc beschrieben.

25 Beispiel 3: Isolierung des lysC gens aus dem C. glutamicum Stamm LU1479

Im ersten Schritt der Stammkonstruktion soll ein allelischer Austausch des lysC Wildtypgens, kodierend für das Enzym Aspartatkinase, in C. glutamicum ATCC13032, im folgenden LU1479 genannt, durchgeführt werden. Dabei soll im LysC Gen ein Nukleotidaustausch durchgeführt werden, so dass im resultierenden Protein die Aminosäure Thr an der Position 311 durch die Aminosäure Ile ausgetauscht ist.

Ausgehend von der chromosomalen DNA aus LU1479 als Template für eine PCR Reaktion

wurde mit den Oligonukleotidprimem SEQ ID NO:65 und SEQ ID NO:66 lysC mit Hilfe des Pfu-Turbo PCR Systems (Stratagene USA) nach Angaben des Herstellers amplifiziert. Chromosomale DNA aus C. glutamicum ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert. Das amplifizierte Fragment wird an seinem 5'-Ende von einem Sall Restriktionsschnitt und an seinem 3'-Ende von einem Mlul Restriktionsschnitt flankiert. Vor der Klonierung wurde das amplifizierte Fragment durch diese beiden Restriktionsenzyme verdaut und mit GFX™PCR, DNA and Gel Band Punfication Kit (Amersham Pharmacia, Freiburg) aufgereinigt.

10 SEQ ID NO:65

5'-GAGAGAGAGACGCGTCCCAGTGGCTGAGACGCATC -3'

SEQ ID NO:66

5'-CTCTCTCTGTCGACGAATTCAATCTTACGGCCTG-3'

15

20

25

Das erhaltenen Polynukleotid wurde über die Sall und Mlul Restriktionsschnitte in pCLIK5 MCS integrativ SacB (im folgenden pCIS genannt; SEQ ID NO: 64 aus Beispiel 2) kloniert und in E.coli XL-1 blue transformiert. Eine Selektion auf Plasmid-tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml)-haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht. Das Plasmid wurden isoliert und durch Sequenzierung die erwartete Nukleotidsequenz bestätigt. Die Präparation der Plasmid-DNA wurde nach Methoden und mit Materialien der Firma Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet. Das erhaltene Plasmid pCIS lysC ist als SEQ ID NO:77 aufgeführt.

Die Sequenz SEQ ID NO:77 umfasst die folgenden wesentlichen Teilbereiche:

Position	Art der Se-	Beschreibung
	quenz	
155 – 1420	CDS ¹⁾	lysC
1974 – 2765	CDS	Kanamycin-Resistenz

3032 – 3892	CDS	Replikationsursprung/
(complement) ²⁾	} 	E.coli/Plasmid pMB

¹⁾ Kodierende Sequenz

Beispiel 4: Mutagenese des lysC Gens aus C. glutamicum

Die gerichtete Mutagenese des lysC Gens aus C. glutamicum (Beispiel 3) wurde mit dem QuickChange Kit (Fa. Stratagene/USA) nach Angaben des Herstellers durchgeführt. Die Mutagenese wurde im Plasmid pCIS lysC, SEQ ID NO:77 durchgeführt. Für den Austausch von thr311 nach 311ile mit Hilfe der Quickchange Methode (Stratagene) wurden folgende Oligonukleotidprimer synthetisiert

10

SEQ ID NO:67

5'-CGGCACCACCGACATCATCTTCACCTGCCCTCGTTCCG -3'

SEQ ID NO:68

15 5'-CGGAACGAGGGCAGGTGAAGATGATGTCGGTGGTGCCG -3'

Der Einsatz dieser Oligonukleotidprimer in der Quickchange Reaktion führt in dem lysC Gen zu einem Austausch des Nukleotids in Position 932 (von C nach T) (vgl. SEQ ID NO:75) und im korrespondierenden Enzym zu einem Aminosäuresubstitution in Position 311 (Thr→IIe) (vgl. SEQ ID NO:76). Der resultierende Aminosäureaustausch Thr311IIe im lysC Gen wurde nach Transformation in E.coli XL1-blue und Plasmidpräparation durch Sequenzierung bestätigt. Das Plasmid erhielt die Bezeichnung pCIS lysC thr311iIe und ist als SEQ ID NO:78 aufgeführt.

Die Sequenz SEQ ID NO:78 umfasst die folgenden wesentlichen Teilbereiche:

25

Position	Art der Se-	Beschreibung
	quenz	
155 – 1420	CDS ¹⁾	lysC mutiert
1974 – 2765	CDS	Kanamycin-Resistenz

²⁾ auf Komplementärstrang

WO 03/087386

39

3032 – 3892	CDS	Replikationsursprung/
(complement) ²⁾		E.coli/Plasmid pMB

¹⁾ Kodierende Sequenz

Das Plasmid pCIS lysC thr311ile wurde in C. glutamicum LU1479 mittels Elektroporation wie bei
Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303 beschrieben, transformiert. Modifikationen des Protokolls sind in DE-A-10046870 beschrieben. Die chromosomale Anordnung des lysC-Lokus einzelner Transformanten wurde mit Standardmethoden durch Southemblot und Hybridisierung, wie in Sambrook et al. (1989), Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben, überprüft. Dadurch wurde sichergestellt, dass es sich bei den Transformanten um solche handelt, die das transformierte Plasmid durch homologe Rekombination am lysC-Lokus integriert haben. Nach Wachstum solcher Kolonien über Nacht in Medien, die kein Antibiotikum enthielten, wurden die Zellen auf ein Saccharose-CM-Agarmedium (10% Saccharose) ausplattiert und bei 30°C für 24 Stunden inkubiert.

Da das im Vektor pCIS lysC thr311ile enthaltende sacB Gen Saccharose in ein toxisches Produkt umwandelt, können nur solche Kolonien anwachsen, die das sacB Gen durch einen zweiten homologen Rekombinationsschritt zwischen dem Wildtyp lysC Gen und dem mutierten Gen lysC thr311ile deletiert haben. Während der homologen Rekombination kann entweder das Wildtyp Gen oder das mutierte Gen zusammen mit dem sacB Gen deletiert werden. Wenn das sacB Gen zusammen mit dem Wildtyp Gen entfernt wird, resultiert eine mutierte Transformante.

20

25

15

Anwachsende Kolonien wurden gepickt, und auf eine Kanamycin-sensitiven Phänotyp hin untersucht. Klone mit deletiertem SacB Gen müssen gleichzeitg Kanamycin-sensitives Wachstumsverhalten zeigen. Solche Kan-sensitiven Klone wurde im einem Schüttelkolben auf ihre Lysin-Produktivität hin untersucht (siehe Beispiel 6). Zum Vergleich wurde der nichtbehandelte Stamm LU1479 angezogen. Klone mit einer gegenüber der Kontrolle erhöhten Lysin-Produktion wurden selektiert, chromosomale DNA wurde gewonnen und der entsprechende Bereich des lysC Gens wurde durch eine PCR-Reaktion amplifiziert und sequenziert. Ein solcher Klon mit der Eigenschaft erhöhter Lysin-Synthese und nachgewiesener Mutation in lysC an der Stelle 932 wurde mit LU1479 lysC 311ile bezeichnet).

²⁾ auf Komplementärstrang

Beispiel 5: Herstellung Ethionin-resistenter C. glutamicum Stämme

Im zweiten Schritt der Stammkonstruktion wurde der erhaltene Stamm LU1479 lysC 311ile (Beispiel 4) behandelt, um eine Ethionin-Resistenz (Kase, H. Nakayama K.Agr. Biol. Chem. 39 153-106 1975 L-methionine production by methionine analog-resistant mutants of Corynebacterium glutamicum) zu induzieren: Eine Übernachtkultur in BHI-Medium (Difco) wurde in Citratpuffer (50mM pH 5,5) gewaschen und bei 30°C für 20 min mit N-Methyl-nitrosoguanidin (10mg/ml·in 50mM Citrat pH5,5) behandelt. Nach der Behandlung mit dem chemischen Mutagen N-Methyl-10 nitrosoguanidin wurden die Zellen gewaschen (Citratpuffer 50mM pH 5,5) und auf ein Medium plattiert, das aus folgenden Komponenten, berechnet auf 500ml, zusammengesetzt war: 10g (NH₄)₂SO₄, 0.5g KH₂PO₄, 0.5g K₂HPO₄, 0.125g MgSO₄-7H₂O, 21g MOPS, 50mg CaCl₂, 15mg Proteokatechuat, 0,5mg Biotin, 1mg Thiamin, 5g/I D,L-Ethionin (Sigma Chemicals Deutschland), pH 7,0. Außerdem enthielt das Medium 0.5ml einer Spurensalzlösung aus: 10g/l FeSO4·7H2O, 1g/I MnSO₄*H₂O₂, 0.1g/I ZnSO₄*7H₂O₂, 0.02g/I CuSO₄, 0.002g/I NiCl₂*6H₂O₂, Alle Salze wurden in 15 0,1M HCl gelöst. Das fertig zusammengestellte Medium wurde sterilfiltriert und nach Zugabe von 40ml steriler 50% Glucoselösung, mit flüssigem sterilem Agar in einer Endkonzentration von 1,5% Agar versetzt und in Kulturschalen ausgegossen.

- 20 Auf Platten mit dem beschriebenen Medium wurden mutagenisierte Zellen aufgebracht und 3-7 Tage bei 30°C inkubiert. Erhaltene Klone wurden isoliert, mindestens einmal auf dem Selektionsmedium vereinzelt und dann auf ihre Methionin-Produktivität in einem Schüttelkolben in Medium II untersucht (siehe Beispiel 6
- 25 Beispiel 6: Herstellung von Methionin mit dem Stamm LU1479 lysC 311ile ET-16.

Die in Beispiel 5 hergestellten Stämme wurden auf einer Agar-Platte mit CM-Medium für 2 Tag bei 30°C angezogen.

CM-Agar:

30 10,0 g/l D-Glucose, 2,5 g/l NaCl, 2,0 g/l Harnstoff, 10,0 g/l Bacto Pepton (Difco), 5,0 g/l Yeast Extract (Difco), 5,0 g/l Beef Extract (Difco), 22,0 g/l Agar (Difco), autoklaviert (20 min., 121°C)

Anschließend wurden die Zellen von der Platte abgekratzt und in Saline resuspendiert. Für die Hauptkultur wurden 10 ml Medium II und 0,5 g autoklaviertes CaCO₃ (Riedel de Haen) in einem 100 ml Erlenmeyerkolben mit der Zellsuspension bis zu einer OD600nm von 1,5 beimpft und für 72h auf einem Orbitalschüttler mit 200 Upm bei 30°C inkubiert.

Medium II:

5

25

30

40g/l Saccharose 60g/l Melasse (auf 100% Zuckergehalt berechnet) 10 10g/I (NH₄)₂SO₄ 0.4g/i MgSO₄*7H₂O 0.6g/IKH₂PO₄ 0.3mg/l Thiamin*HCI 1mg/l Biotin (aus einer 1 mg/ml steril filtrierten Stammlösung die mit NH₄OH auf pH 15 8,0 eingestellt wurde) 2mg/l FeSO₄

2mg/l MnSO₄
mit NH₄OH auf pH 7,8 eingestellt, autoklaviert (121°C, 20 min). Zusätzlich wird Vitamin B12
(Hydroxycobalamin Sigma Chemicals) aus einer Stammlösung (200 μg/ml, stenl filtriert) bis zu

20 einer Endkonzentration von 100 μg/l zugegeben

Gebildetes Methionin, sowie andere Aminosäuren in der Kulturbrühe wurde mit Hilfe der Aminosäuresäure-Bestimmungsmethode von Agilent auf einer Agilent 1100 Series LC System HPLC. Eine Derivatisierung vor der Säulentrennung mit Ortho-Phthalaldehyd erlaubte die Quantifizierung der gebildeten Aminosäuren. Die Auftrennung des Aminosäuregemisch fand auf einer Hypersil AA-Säule (Agilent) statt.

Solche Klone wurden isoliert, deren Methionin-Produktivität mindestens doppelt so hoch war, wie die des Ausgangsstamm LU1479 lysC 311ile. Ein solcher Klon wurde für die weiteren Versuche eingesetzt und bekam die Bezeichnung LU1479 lysC 311ile ET-16.

Beisplei 7: Klonierung von metH aus Streptomyces coelicolor und Klonierung in das Plasmid pC

Phsdh metH_Sc

a) Chromosomale DNA wurde aus *Streptomyces coelicolor* Stamm ATCC BAA-471 (von der American Type Strain Culture Collection, (ATCC) Atlanta, USA, unter der Bestellnummer BAA-471D erhältlich) isoliert. Chromosomale DNA aus C. glutamicum ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert.

Mit den Oligonukleotidprimer SEQ ID NO:69 und SEQ ID NO:70, der chromosomalen DNA aus
 C. glutamicum als Template und Pfu Turbo Polymerase (Fa. Stratagene) wurde mit Hilfe der Polymerase-Kettenreaktion (PCR) nach Standardmethoden, wie Innis et al. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press, ein DNA Fragment von ca. 180 Basenpaaren aus dem nichtkodierenden 5'-Bereich (Promotorregion) der Homoserindehydrogenase (HsDH) amplifiziert. Das amplifizierte Fragment ist an seinem 5'-Ende von einer Xhol-Restriktionsschnittstelle und am 3'-Ende von einem über das Oligo eingeführten zu metH aus Streptomyces coelicolor homologen Bereich flankiert.

. **SEQ ID NO:69**

`5'-GAGACTCGAGGGAAGGTGAATCGAATTTCGG-3'

20 und

30

SEQ ID NO:70

5'-GTCCCGGGGAGAACGCACGATTCTCCAAAAATAATCGC-3'

Das erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

b) Ausgehend von der chromosomalen DNA aus Streptomyces coelicolor als Template für eine PCR Reaktion wurde mit den Oligonukleotidprimern SEQ ID NO:71 und SEQ ID NO:72 ein Teil von metH mit Hilfe des GC-RICH PCR Systems (Roche Diagnostics, Mannheim) nach Angaben des Herstellers amplifiziert. Das amplifizierte Fragment ist an seinem 5'-Ende von einem über das Oligo eingeführten, zur Promotorregion von HsDH aus C. glutamicum homologen Bereich flankiert.

SEQ ID NO:71
5'-GAATCGTGCGTTCTCCCCGGGAC -3'
und
SEQ ID NO:72
5'-GTAGTTGACCGAGTTGATCACC -3'

Das ca. 1,4 kb große erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

c) In einer weiteren PCR Reaktion wurden die beiden oben erhaltenen Fragmente gemeinsam als Template eingesetzt. Durch die mit dem Oligonukleotidprimern SEQ ID NO:71 und SEQ ID NO:70 eingebrachten, zu dem jeweils anderen Fragment homologen Bereichen, kommt es im Zuge der PCR-Reaktion zu einer Anlagerung beider Fragmente aneinander und einer Verlängerung zu einem durchgehenden DNA-Strang durch die eingesetzte Polymerase. Die Standardmethode wurde dahingehend modifiziert, dass die verwendeten Oligonukleotidprimer SEQ ID NO:69 und SEQ ID NO:72 erst mit Beginn des 2. Zyklus dem Reaktionsansatz zugegeben wurden.

Das amplifizierte DNA Fragment von ungefähr 1,6 kb wurde mit dem GFXTMPCR, DNA and Gel Band Purification Kit nach Angaben des Herstellers gereinigt. Im Anschluss daran wurde es mit den Restriktionsenzymen Xhol und Notl (Roche Diagnostics, Mannheim) gespalten und gelelektrophoretisch aufgetrennt. Anschließend wurde das ca. 1,6 kb große DNA Fragment mit GFXTMPCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aus der Agarose aufgereinigt.

25

30

20

d) Der noch fehlende 3'-Bereich von metH wurde ausgehend von der chromosomalen DNA aus Streptomyces coelicolor als Template mit den Oligonukleotidprimem SEQ ID NO:73 und SEQ ID NO:74 mit Hilfe des GC-RICH PCR Systems (Roche Diagnostics, Mannheim) nach Angaben des Herstellers amplifiziert. Das amplifizierte Fragment ist an seinem 3'-Ende von einer über das Oligo eingeführten EcoRV-Restriktionsschnittstelle flankiert.

SEQ ID NO:73 5'-CCGGCCTGGAGAAGCTCG-3'

und

15

20

25

SEQ ID NO:74

5'-GAGAGATATCCCTCAGCGGGCGTTGAAG-3'

- Das erhaltene ca. 2,2 kb große DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Im Anschluß daran wurde es mit den Restriktionsenzymen Notl und EcoRV (Roche Diagnostics, Mannheim) gespalten und gelelektrophoretisch aufgetrennt. Anschließend wurde das ca. 2,2 kb große DNA Fragment mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aus der Agarose aufgereinigt.
 - e) Der Vektor pClik5MCS SEQ ID NO:63 (Beispiel 1) wurde mit den Restriktionsenzymen Xhol und EcoRV (Roche Diagnostics, Mannheim) geschnitten und ein 5 kb großes Fragment nach elektrophoretischer Auftrennung mit GFXTMPCR, DNA and Gel Band Purification Kit isoliert.

Das Vektorfragment wurde zusammen mit den beiden geschnittenen und aufgereinigten PCR-Fragmenten mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid-tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Präparation der Plasmid DNA wurde nach Methoden und mit Materialien der Fa. Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pC Phsdh metH_Sc (Streptomyces coelicolor) ist als SEQ ID NO:79 aufgeführt.

Die Sequenz SEQ ID NO:79 umfasst die folgenden wesentlichen Teilbereiche:

Position	Art der Sequenz	Beschreibung
6 - 155	Promotor	HsDH
156 – 3752	CDS ¹⁾	MetH S. coelicolor
4153 - 4944	CDS	Kanamycin-Resistenz
5211 - 6071 (complement) ²⁾	CDS	Replikationsursprung/ E.coli/Plasmid pMB

¹⁾ Kodierende Sequenz

Beispiel 8: Transformation des Stammes LU1479 lysC 311ile ET-16 mit dem Plasmid pC Phsdh metH_Sc

Der Stamm LU1479 lysC 311ile ET-16 (Beispiel 5) wurde mit dem Plasmid pC Phsdh metH_Sc (Beispiel 7) nach der beschriebenen Methode (Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303) transformiert. Die Transformationsmischung wurde auf CM-Platten plattiert, die zusätzlich 20mg/l Kanamycin enthielten, um eine Selektion auf Plasmid-haltige Zellen zu erreichen. Erhaltene Kan-resistente Klone wurden gepickt und vereinzelt. Die Methionin-Produktivität der Klone wurde in einem Schüttelkolbenversuch (s. Beispiel 6) untersucht. Der Stamm LU1479 lysC 311ile ET-16 pC Phsdh metH_Sc produzierte im Vergleich zu LU1479 lysC 311ile ET-16 signifikant mehr Methionin.

²⁾ auf Komplementärstrang

Patentansprüche

- Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen
 Feinchemikalie, welches folgende Schritte umfasst:
 - a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit Methionin-Synthase (metF) – Aktivität kodiert;
 - b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und
 - c) Isolieren der schwefelhaltigen Feinchemikalie.
- Verfahren nach Anspruch 1, wobei die schwefelhaltige Feinchemikalie L-Methlonin
 umfasst.
 - Verfahren nach einem der vorhergehenden Ansprüche, wobei sich die heterologe metFkodierende Nukleotidsequenz zur metF-kodierenden Sequenz aus Corynebacterium glutamicum ATCC 13032 eine Sequenzhomologie vom weniger als 100% aufweist.

20

10

4. Verfahren nach Anspruch 3, wobei die metF-kodierende Sequenz aus einem der folgenden Organismen abgeleitet ist:

Organimsus	Stammsammlung
Corynebacterium diphteriae	ATCC 14779
Streptomyces lividans	ATCC 19844
Streptomyces coelicolor	ATCC 10147
Aquifex aeolicus	DSM 6858
Burkholderia cepacia	ATCC 25416
Nitrosomonas europaea	ATCC 19718
Pseudomonas aeruginosa	ATCC 17933
Xylella fastidiosa	ATCC 35881
Pseodomonas fluorescens	ATCC 13525
Schizosaccharomyces pombe	ATCC 24969
Saccharomyces cerevisiae	ATCC 10751
Erwinia carotovora	ATCC 15713
Klebsiella pneumoniae	ATCC 700721
Salmonella typhi	ATCC 12839
Salmonella typhimurium	ATCC 15277
Escherichia coli K12	ATCC55151

10

Vibrio cholerae	ATCC 39315
Haemophilus influenzae	ATCC 51907
Caulobacter crescentus	ATCC 19089
Actinobacillus actinomycetemcomitans	ATCC 33384
Neisseria meningitis	ATCC 6253
Rhodobacter capsulatus	ATCC 11166
Campylobacter jejuni	ATCC 33560
Lactococcus lactis	ATCC 7962
Prochlorococcus marinus	PCC7118
Bacillus stearothermophilus	ATCC 12980

- Verfahren nach einem der vorhergehenden Ansprüche, wobei die metF-kodierende Sequenz eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metF-Aktivität kodiert, umfasst.
- Verfahren nach einem der vorhergehenden Ansprüche, wobei die metF-kodierende Sequenz für ein Protein mit metF-Aktivität kodiert, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metF-Aktivität steht, umfasst.
- Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metFSequenz eine in coryneformen Bakterien replizierbare oder eine stabil in das
 Chromosom intregrierte DNA oder eine RNA ist.
 - 8. Verfahren gemäß Anspruch 7, wobei man
 - einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metF-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
 - b) einen Stamm einsetzt, in dem die kodierende metF-Sequenz in das Chromosom des Bakteriums integriert wurde
- Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metF-Sequenz überexprimiert wird.

10

25

30

- 10. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie verstärkt ist oder derart mutiert ist, dass es durch Stoffwechselmetabolite nicht in seiner Aktivität beeinflusst wird.
- 11. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.
- 12. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter
 - a) dem für eine Aspartatkinase kodierenden Gen lysC,
- b) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen gap,
 - c) dem für die 3-Phosphoglycerat Kinase kodierenden Gen pgk,
 - d) dem für die Pyruvat Carboxylase kodierenden Gen pyc,
 - e) dem für die Triosephosphat Isomerase kodierenden Gen tpi,
 - f) dem für die Homoserin O-Acetyltransferase kodierenden Gen metA,
- 20 g) dem für die Cystahionin-gamma-Synthase kodierenden Gen metB,
 - h) dem für die Cystahionin-gamma-Lyase kodierenden Gen metC,
 - i) dem für die Serin-Hydroxymethyltransferase kodierenden Gen glyA,
 - j) dem f
 ür die O-Acetylhomoserin-Sulfhydrylase kodierenden Gen metY,
 - k) dem für das metH Gen, das für die Vitamin B12 abhängige Methionin-Synthase kodiert,
 - I) dem für das serC Gen, das für die Phosphoserin-Aminotransferase kodiert,
 - m) dem serB Gen, das für die Phosphoserin-Phosphatase kodiert,
 - n) dem cysE Gen, das für die Serine Acetyl-Transferase kodiert, und
 - o) dem hom Gen, das eine Homoserin-Dehydrogenase kodiert,

Oberexprimiert oder so mutiert ist, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden.

- 13. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneformen Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter dem für die Homoserine-Kinase kodierenden Gen thrB, a) b) dem für die Threonin Dehydratase kodierenden Gen ilvA. 5 c) dem für die Threonin Synthase kodierenden Gen thrC d) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck, e) f) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgl. g) dem für die Pyruvat-Oxidase kodierenden Gen poxB. 10 h) dem für die Dihydrodipicolinat Synthase kodiemden Gen dapA. i) dem für die Dihydrodipicolinat Reduktase kodiernden Gen dapB; oder j) dem für die Diaminopicolinat Decarboxylase kodiernden Gen
- durch Veränderung der Expressionsrate oder durch Einführung einer gezielten Mutation abschwächt ist.
 - 14. Verfahren gemäß einem oder mehreren der vorhergehenden Ansprüche, wobei man Mikroorganismen der Art Corynebacterium glutamicum einsetzt.
- 20 15. Verfahren zur Herstellung eines L-Methionin haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst
 - a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
 - b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
- 25 c) Entfermung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
 - d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.
- 30 16. Verfahren gemäß Anspruch 15, wobei man Mikroorganismen gemäß der Definition in einem der Ansprüche 1 bis 14 einsetzt.

1 SEQUENZPROTOKOLL

< T	10>	BASF	AKE	ieng	esel.	lsch	aft									
<1	20>	MetH											•	•		
<1	30>	M/43	120											•		
<1	40>															
<14	41>								••				٠,			
<10	50> '	79														
	10> :															
<21	l1> 3	3597														
<21	l2> I	ANC														
<21	13> \$	Stre	ptomy	усев	coel	licol	lor									
<22 <22	!0> !1> (cos														
		(1)	(350	941												
		RSX14		, • ,												
	0> 1															
Val	negi	. Cor	. CCC	cgg	gac	: gcc	: cca	cga	cgg	gcg	gca	ccg	ggc	aga	ggc	48
vai	. ALC) Se1	PIC	AIG	Asp	val	Pro	Arg			Ala	Pro	Gly	Arg	Gly	
				5					10					15		
aaa	gcc	gac	ago	cgt	cgc	ato	cta	999	agc	cct	ttc	atg	gcc	tcg	tcg	96
Lys	Ala	ı Asp	Ser	: Arg	Arg	Ile	Leu	Gly	Ser	Pro	Phe	Met	Ala	Ser	Ser	
			20)				25			·		30			
cca	tcc	acc	ccg	ccc	gcc	gac	acc	cgc	acc	cgc	gtg	tcc	qcc	ctc	cga	144
Pro	Ser	Thr	Pro	Pro	Ala	Asp	Thr	Arg	Thr	Arg	Val	Ser	Ala	Leu	Arg	
		35	,				40					45			5	•
gag	gcc	ctc	gcc	acc	cqc	ata	ata	atc	acc	gac	gge	acc	ata	666	acc	100
Glu	Ala	Leu	Ala	Thr	Ara	Val	Val	Val	Ala	Asn	29c	Ala	Mot	990	acc mb-	192
	50				3	55		• • • •	ALU	vob	60	AIG	Mec	GIY	inr	
ato	ctc	Can			225		•									
Met	Leu	Gln	Ala	cag	Aan	Dro	acg	ctg	gac	gac	ttc	cag	cag	ctc	gaa	240
65	Deu	0111	710	Gln	70	PIO	inr	Leu	Авр		Phe	Gln	Gln	Leu		
										75					80	
999	tgc	aac	gag	gtc	ctg	aac	ctc	acc	cgg	ccc	gac	atc	gtc	cgc	tcg	288
GIY	Сув	Asn	Glu	Val	Leu	Asn	Leu	Thr	Arg	Pro	Авр	Ile	Val	Arg	Ser	
				85					90		_			95		
gtg	cac	gag	gag	tac	ttc	gcg	gcc	ggc	gtc	gac	tac	atc	gag	acc	aac	336
Val	His	Glu	Glu	Tyr	Phe	Ala	Ala	Gly	Val	Asp	Cvs	Val	Glu	Thr	Asn	330
			100					105	•	•	•		110			
366	++-			25-												
acc	Dha	ggc	gcc	aac	cac	tcc	gcc	ctg	ggc	gag	tac	gac	atc	CCC	gag	384
ini	rne	GIA	Ala	Asn	H18	Ser	Ala	Leu	Gly	Glu	Tyr	qaA	Ile	Pro	Glu	
		115					120					125				
cgc	gtc	cac	gaa	ctg	tcc	qaq	qcc	gac	acc	cac	atc	000			a+-	430
Arg	Val	His	Glu	Leu	Ser	Glu	Ala	Glv	Ala	720	Val	ACC.	rac	gag	gtc	432
-	130					135		1		9	140	u14	wid	oru	AGT	
				_												
gcc	gac	gag	ttc	ggc	gcc	cgc	gac	ggc	cgg	cag	cgc	tgg	gtg	ctg	ggc	480

										2						
Ala 145		Glu	Phe	Gly	Ala 150		Asp	Gly	Arg	Gln 155	_	Trp	Val	Leu	Gly 160	
tcc Ser	atg Met	ggc Gly	ccc Pro	ggc Gly 165	acc Thr	aag Lys	ctc Leu	ccc Pro	acc Thr 170	Leu	ggc	cac His	gcc Ala	ccg Pro 175	tac Tyr	528
acc Thr	gtc Val	ctg Leu	cgc Arg 180	Asp	gcc Ala	tac Tyr	cag Gln	cgc Arg 185	aac Asn	gcc Ala	gag Glu	gga Gly	ctg Leu 190	gtc Val	gcg Ala	576
ggc Gly	Gly	gcg Ala 195	gac	gca Ala	ctg Leu	ctg Leu	gtg Val 200	gag Glu	acc	acg Thr	cag Gln	gac Asp 205	ctg Leu	ctc Leu	cag Gln	624
acc Thr	aag Lys 210	gcc Ala	tcg Ser	gtg Val	ctc Leu	ggc Gly 215	gcc Ala	cgg Arg	cgc Arg	gcc Ala	ctg Leu 220	yab	gtc Val	ctc Leu	ggc Gly	672
ctc Leu 225	gac Asp	ctg Leu	ccg Pro	ctc Leu	atc Ile 230	gtg Val	tcc Ser	gtc Val	acc Thr	gtc Val 235	gag Glu	acc Thr	acc Thr	ggc	acc Thr 240	720
Met	Leu	Leu	Gly	Ser 245	Glu	Ile	Gly	Ala	Ala 250	Leu	Thr	Ala	Leu	Glu 255	Pro	768
Leu	Gly	atc Ile	Asp 260	Met	Ile	Gly	Leu	Asn 265	Сув	Ala	Thr	Gly	Pro 270	Ala	Glu	816
Met	Ser	gag Glu 275	His	Leu	Arg	Tyr	Leu 280	Ala	Arg	His	Ser	Arg 285	Ile	Pro	Leu	864
Thr	Сув 290	atg Met	Pro	Asn	Ala	Gly 295	Leu	Pro	Val	Leu	Gly 300	Lys	Asp	Gly	Ala	912
His 305	Tyr	ccg Pro	Leu	Thr	Ala 310	Pro	Glu	Leu	Ala	Asp 315	Ala	His	Glu	Thr	Phe 320	960
Val	Arg	gag Glu	Tyr	Gly 325	Leu	Ser	Leu	Val	Gly 330	Gly	Сув	Сув	Gly	Thr 335	Thr	1008
Pro	Glu	cac His	Leu 340	Arg	Gln	Val	Val	Glu 345	Arg	Val	Arg	Asp	Thr 350	Ala	Pro	1056
Thr	Ala	cgc Arg 355	Asp	Pro	Arg	Pro	Glu 360	Pro	Gly	Ala	Ala	Ser 365	Leu	Tyr	Gln	1104
Thr	Val 370	ccc Pro	Phe	Arg	Gln	Asp 375	Thr	Ser	Tyr	Leu	Ala 380	Ile	Gly	Glu	Arg	1152
acc Thr 385	aac Asn	gcc Ala	aac Asn	Gly	tcc Ser 390	aag Lys	aag Lys	ttc Phe	cgc Arg	gag Glu 395	gcc Ala	atg Met	ctg Leu	gac Asp	ggc Gly 400	1200

cgc Arg	tgg Trp	Asp gac	gac Asp	tgc Cys 405	Val	gag Glu	atg Met	gcc Ala	cgc Arg 410	Asp	cag Gln	atc Ile	cgc Arg	gaa Glu 415	Gly	1248
gcg Ala	cac His	atg Met	ctc Leu 420	Asp	ctc Leu	tgc Cys	gtc Val	gac Asp 425	tac Tyr	gtc Val	ggc Gly	cgg Arg	gac Asp 430	Gly	gtc Val	1296
gcc Ala	gac Asp	atg Met 435	gag Glu	gaa Glu	ctg Leu	gcc Ala	ggc Gly 440	cgg Arg	_ttc Phe	gcc Ala	acc Thr	gcc Ala 445	tcc Ser	acg Thr	ctg Leu	1344
ccg Pro	atc Ile 450	gtc Val	ctc Leu	Asp	tcc Ser	acc Thr 455	gag Glu	gtc Val	gac Asp	gtc Val	atc Ile 460	cgg Arg	gcc Ala	ggc	ctg Leu	1392
gag Glu 465	aag Lys	ctc Leu	ggc	ggc	cgc Arg 470	gcg Ala	gtg Val	atc Ile	aac Asn	tcg Ser 475	gtc Val	aac Asn	tac Tyr	gag Glu	gac Asp 480	1440
Gly	Ala	Gly	Pro	Glu 485	Ser	cgg	Phe	Ala	Arg 490	Val	Thr	Lys	Leu	Ala 495	Arg	1488
Glu	His	Gly	Ala 500	Ala	Leu	atc Ile	Ala	Leu 505	Thr	Ile	Asp	Glu	Val 510	Gly	Gln	1536
Ala	Arg	Thr 515	Ala	Glu	Lys	aag Lys	Val 520	Glu	Ile	Ala	Glu	Arg 525	Leu	Ile	Asp	1584
Ąap	Leu 530	Thr	Gly	Asn	Trp	ggc Gly 535	Ile	His	Glu	Ser	Авр 540	Ile	Leu	Val	Asp	1632
Сув 545	Leu	Thr	Phe	Thr	Ile 550	tgc Cys	Thr	Gly	Gln	Glu 555	Glu	Ser	Arg	Lys	Авр 560	1680
Gly	Leu	Ala	Thr	Ile 565	Glu	ggc Gly	Ile	Arg	Glu 570	Leu	Lys	Arg	Arg	His 575	Pro	1728
Asp	Val	Gln	Thr 580	Thr	Leu	ggc	Leu	Ser 585	Asn	Ile	Ser	Phe	Gly 590	Leu	Asn	1776
Pro	Ala	Ala 595	Arg	Ile	Leu	ctc Leu	Asn 600	Ser	Val	Phe	Leu	Asp 605	Glu	Сув	Val	1824
aag Lys	gcc Ala 610	ggc Gly	ctg Leu	gac Asp	Ser	gcc Ala 615	atc Ile	gtg Val	cac His	gcg Ala	agc Ser 620	aag Lye	atc Ile	ctg Leu	ccg Pro	1872
atc Ile 625	gcc Ala	cgc Arg	ttc Phe	Asp	gag Glu 630	gag Glu	cag Gln	gtc Val	acc Thr	acc Thr 635	gcc Ala	ctc Leu	gac Asp	ttg Leu	atc Ile 640	1920
tac	gac	cgc	cgc	cgc	gag	ggc	tac	gac	ccc	ctg	caa	aag	ctc	atg	cag	1968

										4						
Tyr	Asp	Arg	Arg	Arg 645		Gly	Tyr	Asp	Pro 650		Gln	Lys	Leu	Met 655	Gln	
ctc Leu	ttc Phe	gag Glu	ggc 660	Ala	acc	gcc Ala	aag Lys	tcg Ser 665	ctg Leu	aag Lys	gcc Ala	tcc Ser	aag Lys 670	gcc Ala	gag Glu	2016
gaa Glu	ctg Leu	gcc Ala 675	gcc Ala	ctc Leu	ccg Pro	ctg Leu	gag Glu 680	gag Glu	cgc Arg	ctc Leu	aag Lys	cgc Arg 685	cgc Arg	atc Ile	atc Ile	2064
Asp	ggc Gly 690	gag Glu	aag Lys	aac Asn	ggc	ctc Leu 695	gaa Glu	cag Gln	gac Asp	ctc Leu	gac Asp 700	gag Glu	gcc Ala	ctc Leu	cgg Arg	2112
Glu 705	Arg	Pro	Ala	Leu	Glu 710	atc Ile	Val	Asn	Asp	Thr 715	Leu	Leu	Asp	Gly	Met 720	2160
Lys	Val	Val	Gly	Glu 725	Leu	ttc Phe	Gly	Ser	Gly 730	Gln	Met	Gln	Leu	Pro 735	Phe	2208
Val	Leu	Gln	Ser 740	Ala	Glu	gtc Val	Met	Lys 745	Thr	Ala	Val	Ala	His 750	Leu	Glu	2256
Pro	His	Met 755	Glu	Lys	Thr	gac Asp	Asp 760	Asp	Gly	Lys	Gly	Thr 765	Ile	Val	Leu	2304
Ala	Thr 770	Val	Arg	Gly	Asp	gtc Val 775	His	Asp	Ile	Gly	Lys 780	Asn	Leu	Val	Asp	2352
11e 785	Ile	Leu	Ser	Asn	Asn 790	ggc	Tyr	Asn	Val	Val 795	Asn	Leu	Gly	Ile	Eys 800	2400
Gln	Pro	Val	Ser	Ala 805	Ile	ctg Leu	Glu	Ala	Ala 810	Asp	Glu	His	Arg	Ala 815	Asp	2448
gtc Val	atc Ile	ggc Gly	atg Met 820	tcc Ser	ggc Gly	ctc Leu	ctc Leu	gtc Val 825	aag Lys	tcc Ser	acg Thr	gtg Val	atc Ile 830	atg Met	aag Lys	2496
gag Glu	aac Asn	ctg Leu 835	gag Glu	gag Glu	ctg Leu	aac Asn	cag Gln 840	cgc Arg	aag Lys	ctg Leu	gcc Ala	gcc Ala 845	gac Asp	tac Tyr	ccg Pro	2544
gtc Val	atc Ile 850	ctc Leu	ggc Gly	ggc Gly	gcc Ala	gcc Ala 855	ctc Leu	acc Thr	agg Arg	gcc Ala	tac Tyr 860	gtc Val	gaa Glu	cag Gln	gac Asp	2592
ctg Leu 865	cac His	gag Glu	atc Ile	tac Tyr	gac Asp 870	ggc Gly	gag Glu	gtc Val	cgc Arg	tac Tyr 875	gcc Ala	cgc Arg	gac Asp	gcc Ala	ttc Phe 880	2640
gag Glu	ggc Gly	ctg Leu	Arg	ctc Leu 885	atg Met	gac Asp	gcc Ala	ctc Leu	atc Ile 890	ggc Gly	atc Ile	aag Lys	cgc Arg	ggc Gly 895	gtg Val	2688

ccc Pro	ggc Gly	gcc Ala	aag Lys 900	ctg Leu	ccg	gag Glu	ctg Leu	aag Lys 905	cag Gln	cgc	cgg	Val	cgg Arg 910	gcc Ala	gcc Ala	2736
acc Thr	gtc Val	gag Glu 915	atc Ile	gac Asp	gag Glu	cgc Arg	ccc Pro 920	gag Glu	gaa Glu	ggc Gly	cac His	gtc Val 925	cgc Arg	tcc Ser	gac Asp	2784
gtc Val	gcc Ala 930	Thr	gac Asp	aac Asn	ccg Pro	gtc Val 935	ccg Pro	acc Thr	ccg	ccc Pro	ttc Phe 940	cgc Arg	ggc Gly	acc Thr	cgc Arg	2832
gtc Val 945	gtc Val	aag Lys	ggc Gly	atc Ile	cag Gln 950	ctc Leu	aag Lys	gag Glu	tac Tyr	gcc Ala 955	Ser	tgg Trp	ctc Leu	gac Asp	gag Glu 960	2880
ggc Gly	gcc Ala	ctc Leu	ttc Phe	aag Lys 965	ggc Gly	cag Gln	tgg Trp	ggc Gly	ctc Leu 970	aag Lys	cag Gln	gcc Ala	cgc Arg	acc Thr 975	ggc	2928
Glu	Gly	Pro	Ser 980	Tyr	gag Glu	Glu	Leu	Val 985	Glu	Ser	Glu	Gly	Arg 990	Pro	Arg	2976
Leu	Arg	Gly 995	Leu	Leu	gac Asp	Arg 1	Leu 1000	Gln	Thr	Asp	Asn	Leu 1005	Leu	Glu	Ala	3024
Ala 1	Val .010	Val	Tyr	Gly		Phe 015	Pro	Сув	Val	Ser	Lys 1020	Asp	qaA	Asp	Leu	3072
Ile 1025	Val	Leu	Asp	Asp	gac Asp 1030	Gly	Asn	Glu	Arg	Thr 1035	Arg	Phe	Thr	Phe	Pro 1040	3120
cgc Arg	cag Gln	cgc Arg	Arg	ggc Gly 045	cgg Arg	cgc Arg	ctg Leu	Сув	ctg Leu .050	gcc Ala	Asp	ttc Phe	Phe	cgc Arg	ccg Pro	3168
gag Glu	gag Glu	Ser	ggc Gly 1060	gag Glu	acc Thr	gac Asp	Val	gtc Val .065	ggc Gly	ttc Phe	cag Gln	Val	gtc Val	acc Thr	gtc Val	3216
ggc Gly	Ser	cgc Arg 1075	atc Ile	ggc Gly	gag Glu	Glu	acg Thr 080	gcc Ala	cgc Arg	atg Met	Phe	gag Glu 085	gcc Ala	aac Asn	gcc Ala	3264
Tyr	cgc Arg 090	gac Asp	tat Tyr	ctc Leu	gag Glu 1	ctg Leu 095	cac His	ggc Gly	ctg Leu	Ser	gtg Val .100	cag Gln	ctc Leu	gcc Ala	gag Glu	3312
gcc Ala 1105	Leu	gcc Ala	gag Glu	tac Tyr	tgg Trp 1110	cac His	gcg Ala	cgc Arg	gtg Val	cgc Arg 1115	Ser	gaa Glu	ctc Leu	ggc	ttc Phe 1120	3360
gcc Ala	G1y 999	gag Glu	Asp	ccg Pro 125	gcc Ala	gag Glu	atg Met	Glu	gac Asp 130	atg Met	ttc Phe	gcc Ala	Leu	aag Lys 135	tac Tyr	3408

WO 03/087386 PCT/EP03/04010

cgg ggt gcc cgc ttc tcc ctc ggc tac ggc gcc tgc ccc gac ctg gag Arg Gly Ala Arg Phe Ser Leu Gly Tyr Gly Ala Cys Pro Asp Leu Glu 1140 1145 gac ege gee aag ate gee gee etg etg gag eee gag ege ate gge gte Asp Arg Ala Lys Ile Ala Ala Leu Leu Glu Pro Glu Arg Ile Gly Val 1160 1165 cac cta tcc gag gag ttc cag ctc cac ccc gag cag tcc acc gac gcc 3552 His Leu Ser Glu Glu Phe Gln Leu His Pro Glu Gln Ser Thr Asp Ala 1170 1175 atc gtc atc cac cac ccg gag gcc aag tac ttc aac gcc cgc 3594 Ile Val Ile His His Pro Glu Ala Lys Tyr Phe Asn Ala Arg 1190 1195 tga 3597 <210> 2

<211> 2 <211> 1198 <212> PRT <213> Streptomyces coelicolor

- -

<400> 2

Val Arg Ser Pro Arg Asp Val Pro Arg Arg Ala Ala Pro Gly Arg Gly

1 10 15

Lys Ala Asp Ser Arg Arg Ile Leu Gly Ser Pro Phe Met Ala Ser Ser 20 25 30

Pro Ser Thr Pro Pro Ala Asp Thr Arg Thr Arg Val Ser Ala Leu Arg 35 40 45

Glu Ala Leu Ala Thr Arg Val Val Ala Asp Gly Ala Met Gly Thr
50 60

Met Leu Gln Ala Gln Asn Pro Thr Leu Asp Asp Phe Gln Gln Leu Glu 65 70 75 80

Gly Cys Asn Glu Val Leu Asn Leu Thr Arg Pro Asp Ile Val Arg Ser 85 90 95

Val His Glu Glu Tyr Phe Ala Ala Gly Val Asp Cys Val Glu Thr Asn 100 105 110

Thr Phe Gly Ala Asn His Ser Ala Leu Gly Glu Tyr Asp Ile Pro Glu 115 120 125

Arg Val His Glu Leu Ser Glu Ala Gly Ala Arg Val Ala Arg Glu Val 130 135 140

Ala Asp Glu Phe Gly Ala Arg Asp Gly Arg Gln Arg Trp Val Leu Gly
145 150 155 160

Ser Met Gly Pro Gly Thr Lys Leu Pro Thr Leu Gly His Ala Pro Tyr
165 170 175

Thr Val Leu Arg Asp Ala Tyr Gln Arg Asn Ala Glu Gly Leu Val Ala 180 185 190

Gly Gly Ala Asp Ala Leu Leu Val Glu Thr Thr Gln Asp Leu Leu Gln

195

205

Thr Lys Ala Ser Val Leu Gly Ala Arg Arg Ala Leu Asp Val Leu Gly 210 220

200

Leu Asp Leu Pro Leu Ile Val Ser Val Thr Val Glu Thr Thr Gly Thr 225 230 235 240

Met Leu Gly Ser Glu Ile Gly Ala Ala Leu Thr Ala Leu Glu Pro 245 250 255

Leu Gly Ile Asp Met Ile Gly Leu Asn Cys Ala Thr Gly Pro Ala Glu 260 265 270

Met Ser Glu His Leu Arg Tyr Leu Ala Arg His Ser Arg Ile Pro Leu 275 280 285

Thr Cys Met Pro Asn Ala Gly Leu Pro Val Leu Gly Lys Asp Gly Ala 290 295 300

His Tyr Pro Leu Thr Ala Pro Glu Leu Ala Asp Ala His Glu Thr Phe 305 310 315 320

Val Arg Glu Tyr Gly Leu Ser Leu Val Gly Gly Cys Cys Gly Thr Thr 325 330 335

Pro Glu His Leu Arg Gln Val Val Glu Arg Val Arg Asp Thr Ala Pro 340 345 350

Thr Ala Arg Asp Pro Arg Pro Glu Pro Gly Ala Ala Ser Leu Tyr Gln 355 360 365

Thr Val Pro Phe Arg Gln Asp Thr Ser Tyr Leu Ala Ile Gly Glu Arg 370 375 380

Thr Asn Ala Asn Gly Ser Lys Lys Phe Arg Glu Ala Met Leu Asp Gly 385 390 395 400

Arg Trp Asp Asp Cys Val Glu Met Ala Arg Asp Gln Ile Arg Glu Gly
405 410 415

Ala His Met Leu Asp Leu Cys Val Asp Tyr Val Gly Arg Asp Gly Val
420 430

Ala Asp Met Glu Glu Leu Ala Gly Arg Phe Ala Thr Ala Ser Thr Leu 435 440 445

Pro Ile Val Leu Asp Ser Thr Glu Val Asp Val Ile Arg Ala Gly Leu
450 460

Glu Lys Leu Gly Gly Arg Ala Val Ile Asn Ser Val Asn Tyr Glu Asp 465 470 475 480

Gly Ala Gly Pro Glu Ser Arg Phe Ala Arg Val Thr Lys Leu Ala Arg 485 490 495

Glu His Gly Ala Ala Leu Ile Ala Leu Thr Ile Asp Glu Val Gly Gln 500 505 510

Ala Arg Thr Ala Glu Lys Lys Val Glu Ile Ala Glu Arg Leu Ile Asp 515 520 525

- Asp Leu Thr Gly Asn Trp Gly Ile His Glu Ser Asp Ile Leu Val Asp 530 535 540
- Cys Leu Thr Phe Thr Ile Cys Thr Gly Gln Glu Glu Ser Arg Lys Asp 550 555 560
- Gly Leu Ala Thr Ile Glu Gly Ile Arg Glu Leu Lys Arg Arg His Pro
 565 570 575
- Asp Val Gln Thr Thr Leu Gly Leu Ser Asn Ile Ser Phe Gly Leu Asn 580 585 590
- Pro Ala Ala Arg Ile Leu Leu Asn Ser Val Phe Leu Asp Glu Cys Val
 595 600 605
- Lys Ala Gly Leu Asp Ser Ala Ile Val His Ala Ser Lys Ile Leu Pro 610 620
- Ile Ala Arg Phe Asp Glu Glu Gln Val Thr Thr Ala Leu Asp Leu Ile 625 630 635 640
- Tyr Asp Arg Arg Glu Gly Tyr Asp Pro Leu Gln Lys Leu Met Gln
 645 650 655
- Leu Phe Glu Gly Ala Thr Ala Lys Ser Leu Lys Ala Ser Lys Ala Glu 660 665 670
- Glu Leu Ala Ala Leu Pro Leu Glu Glu Arg Leu Lys Arg Arg Ile Ile 675 680 685
- Asp Gly Glu Lys Asn Gly Leu Glu Gln Asp Leu Asp Glu Ala Leu Arg 690 695 700
- Glu Arg Pro Ala Leu Glu Ile Val Asn Asp Thr Leu Leu Asp Gly Met 705 710 715 720
- Lys Val Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe 725 730 735
- Val Leu Gln Ser Ala Glu Val Met Lys Thr Ala Val Ala His Leu Glu 740 745 750
- Pro His Met Glu Lys Thr Asp Asp Gly Lys Gly Thr Ile Val Leu 755 760 765
- Ala Thr Val Arg Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Asp
 770 780
- Ile Ile Leu Ser Asn Asn Gly Tyr Asn Val Val Asn Leu Gly Ile Lys
 785 790 795 800
- Gln Pro Val Ser Ala Ile Leu Glu Ala Ala Asp Glu His Arg Ala Asp 805 810 815
- Val Ile Gly Met Ser Gly Leu Leu Val Lys Ser Thr Val Ile Met Lys 820 825 830
- Glu Asn Leu Glu Glu Leu Asn Gln Arg Lys Leu Ala Ala Asp Tyr Pro 835 840 845
- Val Ile Leu Gly Gly Ala Ala Leu Thr Arg Ala Tyr Val Glu Gln Asp 850 855 860

- Leu His Glu Ile Tyr Asp Gly Glu Val Arg Tyr Ala Arg Asp Ala Phe 865 870 875 880
- Glu Gly Leu Arg Leu Met Asp Ala Leu Ile Gly Ile Lys Arg Gly Val 885 890 895
- Pro Gly Ala Lys Leu Pro Glu Leu Lys Gln Arg Arg Val Arg Ala Ala 900 905 910
- Thr Val Glu Ile Asp Glu Arg Pro Glu Glu Gly His Val Arg Ser Asp 915 920 925
- Val Ala Thr Asp Asn Pro Val Pro Thr Pro Pro Phe Arg Gly Thr Arg 930 935 940
- Val Val Lys Gly Ile Gln Leu Lys Glu Tyr Ala Ser Trp Leu Asp Glu 945 950 955 960
- Gly Ala Leu Phe Lys Gly Gln Trp Gly Leu Lys Gln Ala Arg Thr Gly 965 970 975
- Glu Gly Pro Ser Tyr Glu Glu Leu Val Glu Ser Glu Gly Arg Pro Arg 980 985 990
- Leu Arg Gly Leu Leu Asp Arg Leu Gln Thr Asp Asn Leu Leu Glu Ala 995 1000 1005
- Ala Val Val Tyr Gly Tyr Phe Pro Cys Val Ser Lys Asp Asp Asp Leu 1010 1015 1020
- Ile Val Leu Asp Asp Asp Gly Asn Glu Arg Thr Arg Phe Thr Phe Pro 1025 1030 1035 1040
- Arg Gln Arg Arg Gly Arg Arg Leu Cys Leu Ala Asp Phe Phe Arg Pro 1045 1050 1055
- Glu Glu Ser Gly Glu Thr Asp Val Val Gly Phe Gln Val Val Thr Val
 1060 1065 1070
- Gly Ser Arg Ile Gly Glu Glu Thr Ala Arg Met Phe Glu Ala Asn Ala 1075 1080 1085
- Tyr Arg Asp Tyr Leu Glu Leu His Gly Leu Ser Val Gln Leu Ala Glu 1090 1095 1100
- Ala Leu Ala Glu Tyr Trp His Ala Arg Val Arg Ser Glu Leu Gly Phe 1105 1110 1115 1120
- Ala Gly Glu Asp Pro Ala Glu Met Glu Asp Met Phe Ala Leu Lys Tyr 1125 1130 1135
- Arg Gly Ala Arg Phe Ser Leu Gly Tyr Gly Ala Cys Pro Asp Leu Glu 1140 1145 1150
- Asp Arg Ala Lys Ile Ala Ala Leu Leu Glu Pro Glu Arg Ile Gly Val 1155 1160 1165
- His Leu Ser Glu Glu Phe Gln Leu His Pro Glu Gln Ser Thr Asp Ala 1170 1175 1180
- Ile Val Ile His His Pro Glu Ala Lys Tyr Phe Asn Ala Arg

10 1185 1190 1195

<210> 3 <211> 3537 <212> DNA <213> Anabaena sp. <220> <221> CDS <222> (1)..(3534) <223> RAN03790 <400> 3 atg act cat cct ttc ctg aaa cgc ctg cac agt ccg gaa ctt ccg gtt Met Thr His Pro Phe Leu Lys Arg Leu His Ser Pro Glu Leu Pro Val ate gte tte gae ggt gea atg gga aet aac eta caa ace caa aac ete Ile Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Thr Gln Asn Leu 20 25 acg gct gag gat ttc ggc ggt gtg cag tat gaa ggt tgt aac gaa tac Thr Ala Glu Asp Phe Gly Gly Val Gln Tyr Glu Gly Cys Asn Glu Tyr 40 cta gtc cac acc aaa ccc gaa gct gtc gcc aag gtt cac cgc gac ttt Leu Val His Thr Lys Pro Glu Ala Val Ala Lys Val His Arg Asp Phe 50 ctc gct gtg ggt gca gat gtc atc gaa acc gac act ttc ggt gcg aca Leu Ala Val Gly Ala Asp Val Ile Glu Thr Asp Thr Phe Gly Ala Thr 65 70 tcc att gtt ttg gcg gaa tat gac tta gca gac caa aca tat tac ctg 288 Ser Ile Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Thr Tyr Tyr Leu 90 95 aac aag aaa gcc gcc gaa ctg gcg aaa agt gtc gct gct gaa ttt tcc 336 Asn Lys Lys Ala Ala Glu Leu Ala Lys Ser Val Ala Ala Glu Phe Ser 100 105 aca cca gat aaa ccc cgg ttt gtt gct ggt tcc atc ggc ccc aca acc 384 Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr 120 aaa ctt ccc acc ttg gga cat atc gac ttt gac act ctc aaa act tgc 432 Lys Leu Pro Thr Leu Gly His Ile Asp Phe Asp Thr Leu Lys Thr Cys 135 ttt gct gaa caa gca gaa gcg ctg tta gat ggt ggc gtg gat tta ctt Phe Ala Glu Gln Ala Glu Ala Leu Leu Asp Gly Gly Val Asp Leu Leu ttg gtg gag act tgt caa gat gtg ctg caa atc aaa gcg gcg ctg aat 528 Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn ggg ata gaa gac ttt ggc aag aga ggg gaa cgc ata ccc ttg atg Gly Ile Glu Glu Val Phe Gly Lys Arg Gly Glu Arg Ile Pro Leu Met 180 185

										11				•		
gtg Val	g to	c gt r Va 19	1 Th	a at	g ga t Gl	a ago u Se:	c atg r Mei 200	t Gly	g aca	a at	g ttg t Le	g gte u Vai 20	1 G1	t tc y Se	c gaa r Glu	624
ato Ile	2 aa 2 Ası 210	u Al	c gt a Va	c ctg l Le	g ac	a att	≥ Lev	a gaa u Glu	a cci	t tto Pho	e Pro	o Ile	t ga	c at p Il	t ctc e Leu	672
ggt Gly 225	, re(g aa 1 As:	c tg n Cy	t gco s Ala	a Thi	r GI	cca Pro	a gad Asp	ttg Lei	ate 1 Mei 23!	Ly	a cca s Pro	a cat	t at	t aaa e Lys 240	720
tat Tyr	ttg Lev	g gc	t gaa a Glu	a cat 1 His 245	s Sei	g ccg	ttt Phe	gtg Val	gtt Val 250	Se	t tgt c Cys	att Ile	e Pro	t aad As: 25!	c gcg n Ala 5	768
GIÀ	Leu	l Pro	260	ı Asn	ı Val	Gly	Gly	Gln 265	Ala	Hia	з Тух	Arg	270	1 Th:	a cca r Pro	816
MEC.	GIU	275	i Arg	, wet	Ala	Leu	280	His	Phe	· Val	. Glu	Asp 285	Lev	Gly	t gtc y Val	864
GIII	290	116	: сту	GIY	Сув	295	Gly	Thr	Arg	Pro	300	His	Ile	Glr	a caa a Gln	912
305	MIG	GIU	1116	ATA	110 310	Asp	Leu	Lys	Pro	1ys 315	Val	Arg	Gln	Pro	agt Ser 320	960
Deu	GIU	PIO	Ala	325	Ala	ser	Ile	Tyr	Ser 330	Thr	Gln	Pro	Tyr	Glu 335		1008
хар	Vell	ser	340	ren	116	Val	Gly	Glu 345	Arg	Leu	Asn	Ala	Ser 350	Gly	tcc Ser	1056
-,,,	275	355	AL Y	web	ren	ctg Leu	360	ATA	GIA	Asp	Trp	365	Gly	Leu	Val	1104
261	370	AIG	Arg	ser	GIN	gtc Val 375	ГÀв	Glu	Gly	Ala	His 380	Ile	Leu	Asp	Val	1152
385	V	veħ	ıyı	Val	390	cgg Arg	Asp	GIÀ	Val	Arg 395	Asp	Met	His	Glu	Leu 400	1200
V	261	AIG	116	405	Asn	aat Asn	val	Thr	Leu 410	Pro	Leu	Met	Leu	Asp 415	Ser	1248
	GIU	irp	420	гув	Met	gag Glu	Ala	Gly 425	Leu	Lys	Val	Ala	Gly 430	Gly	Lys	1296
gt (Cys)	ttg Leu	ctg Leu	aac Asn	tcc Ser	acc Thr	aac Asn	tac Tyr	gaa Glu	gat Asp	999 999	gaa Glu	cca Pro	cgt Arg	ttc Phe	tta Leu	1344

l	3	5						
---	---	---	--	--	--	--	--	--

		435					440					445				
aaa Lys	gtg Val 450	Leu	gag Glu	ttg Leu	gcg Ala	aag Lys 455	aaa Lys	tat Tyr	ggc Gly	gcg Ala	ggt Gly 460	gtt Val	gtt Val	att Ile	ggc	1392
aca Thr 465	Ile	gac Asp	gaa Glu	gaa Glu	999 Gly 470	atg Met	gcg Ala	cgg Arg	aca Thr	gcc Ala 475	gag Glu	aaa Lys	aag Lys	ttt Phe	caa Gln 480	1440
att Ile	gcc Ala	cag Gln	cgt Arg	gcc Ala 485	tat Tyr	cgt Arg	caa Gln	tcg Ser	gta Val 490	gaa Glu	tat Tyr	Gly 999	att Ile	ccc Pro 495	ccc Pro	1488
aca Thr	gaa Glu	ata Ile	ttc Phe 500	ttt Phe	gat Asp	acc Thr	tta Leu	gct Ala 505	tta Leu	cca Pro	att Ile	tct Ser	acc Thr 510	G1y 999	att Ile	1536
gaa Glu	gaa Glu	gac Asp 515	cgg Arg	gaa Glu	aat Asn	ggc Gly	aag Lys 520	gcg	aca Thr	att Ile	gaa Glu	tca Ser 525	att Ile	agc Ser	cgt Arg	1584
Ile	Arg 530	Lys	gaa Glu	Leu	Pro	Gly 535	Сув	His	Val	Ile	Leu 540	Gly	Val	Ser	Asn	1632
Ile 545	Ser	Phe	ggc Gly	Leu	Asn 550	Ser	Ala	Ser	Arg	Met 555	Val	Leu	Asn	Ser	Val 560	1680
Phe	Leu	His	gaa Glu	Ala 565	Met	Thr	Ala	Gly	Met 570	yeb	Ala	Ala	Ile	Val 575	Ser	1728
Ala	Ser	Lys	att Ile 580	Leu	Pro	Leu	Ser	Lys 585	Ile	Glu	Glu	Arg	His 590	Gln	Glu	1776
Val	Сув	Arg 595	cag Gln	Leu	Ile	Tyr	Asp	Gln	Arg	Lys	Phe	Glu 605	Gly	Asp	Ile	1824
Сув	Ile 610	Tyr	gac Asp	Pro	Leu	Thr 615	Glu	Leu	Thr	Lys	Leu 620	Phe	Glu	Gly	Val	1872
Thr 625	Thr	Lys	cgt Arg	Asn	Lys 630	Gly	Val	Asp	Glu	Ser 635	Leu	Pro	Ile	Glu	Glu 640	1920
cga Arg	ctc Leu	aag Lys	cgt Arg	cac His 645	att Ile	atc Ile	gac Asp	ggc Gly	gaa Glu 650	cgc Arg	att Ile	ggt Gly	tta Leu	gaa Glu 655	gcg Ala	1968
caa Gln	ctg Leu	aca Thr	aaa Lys 660	gcc Ala	tta Leu	gaa Glu	caa Gln	tat Tyr 665	cca Pro	ccc Pro	cta Leu	gaa Glu	att Ile 670	atc Ile	aac Asn	2016
act Thr	ttc Phe	cta Leu 675	cta Leu	gat Asp	G1 y 999	atg Met	aaa Lys 680	gta Val	gtc Val	G] À 888	gaa Glu	ttg Leu 685	ttc Phe	ggt Gly	tca Ser	2064

										13						
G1	a ca y Gl 69	n Me	g ca t Gl	g ct n Le	a cct u Pro	Phe 695	e Vai	t tta l Le	a cag a Glr	j tca 1 Sei	a gce c Ala 700	a Glu	a ace	c ato	aaa Lys	2112
gc; A1: 70:	a Al	g gt a Va	a gc 1 Al	c ta a Ty:	c cta r Lev 710	Glu	ccg Pro	tto Phe	ato Met	gaa Glu 715	ı Lya	a tog Sei	g gaa	a agt u Sei	ggc Gly 720	2160
aa Ası	c aat	t gc	c aa: a Ly:	a gg(s Gl) 729	y Lys	gta Val	att Ile	att lle	gcc Ala 730	Thr	gtg Val	g aas Lys	gge Gl	c gat y As <u>r</u> 735	gtt Val	2208
Cac Hi	c gae	c at	t gg(e Gl) 740	y Lys	a aac s Asn	cta Leu	gta Val	gac Asp 745	Ile	ato Ile	ttg Lev	tco Ser	A81	a Asr	ggc Gly	2256
tac Ty:	aag Lys	g gta 3 Va: 755	LIle	aac Asr	ctg Leu	gga Gly	att Ile 760	Lys	cag Gln	ccg Pro	gtg Val	gaa Glu 765	Asr	ato lle	atc lle	2304
gag	get Ala 770	і Туі	аас Авг	caa Glm	cac His	aaa Lys 775	gct	gat Asp	tgt Cys	att Ile	gcc Ala 780	Met	agt Ser	ggc Gly	ttg Leu	2352
785	i val	. Lys	Ser	Thr	gca Ala 790	Phe	Met	Lys	Glu	Asn 795	Leu	Glu	Val	Phe	Asn 800	2400
GIU	гув	GIY	, 116	805		Pro	Val	Ile	Leu 810	Gly	Gly	Ala	Ala	815	Thr	2448
ccg Pro	ааа Lув	Phe	gtg Val 820	H18	aaa Lys	gat Asp	tgc Cys	caa Gln 825	aat Asn	acc Thr	tac Tyr	aaa Lys	ggt Gly 830	Lys	gtc Val	2496
att Ile	tat Tyr	ggc Gly 835	aaa Lys	gat Asp	gct Ala	ttc Phe	tca Ser 840	yab Gac	ctg Leu	cat His	ttc Phe	atg Met 845	gat Asp	aaa Lys	tta Leu	2544
мет	850	Ala	Lys	Ala	act Thr	Gly 855	Lys	Trp	Asp	Asn	Ser 860	Leu	Gly	Phe	Leu	2592
865	GII	val	Glu	Thr	gag Glu 870	Glu	Thr	Glu	Pro	Thr 875	Asn	His	Lys	Ser	Pro 880	2640
atc Ile	ccc Pro	agt Ser	ccc Pro	caa Gln 885	tcc Ser	cca Pro	gtc Val	ccc Pro	agt Ser 890	ccc Pro	cag Gln	tcc Ser	cca Pro	gtc Val 895	cct Pro	2688
ata Ile	Aap Aap	acc Thr	cga Arg 900	cgt Arg	tcc Ser	gaa Glu	Ala	gta Val 905	gcc Ala	ata Ile	gac Asp	att Ile	ccc Pro 910	cgt Arg	ccc Pro	2736
Int	PIO	915	Pne	тър	gga ; Gly !	Thr (Gln 920	Leu :	Leu	Gln	Pro	Ser 925	qaA	Ile	Ser	2784
tta	gag	gaa	ata	ttc	tgg (cac .	atg	gat	ttg	caa	gcc	ttg	att	gcg	gga	2832

										14						
Leu (Glu 930	Glu	Ile	Phe	Trp	935	Met	: Asp) Lev	Gln	Ala 940	Leu	Ile	Ala	Gly	
Caa (Gln 5 945	tgg Irp	caa Gln	ttc Phe	cgc Arg	aaa Lys 950	Pro	aaa Lys	gaa Glu	caa Gln	tca Ser 955	aag Lys	gaa Glu	gaa Glu	tat Tyr	Gln 960	2880
gct t Ala I	ene :	Leu	Asn	965	Lys	Val	Tyr	Pro	970	. Leu	Glu	Thr	Trp	Lys 975	Gln	2928
cgc a	ie.	116	980	GIu	Asn	Leu	Leu	Н18 985	Pro	Gln	Val	Ile	Tyr 990	Gly	Tyr	2976
ttt c Phe P	ro (сув 995	Gln	Ser	Glu	Gly	Asn 1000	Thr	Leu	Tyr	Val	Tyr 1005	Glu	Thr	Asn	3024
	10 10	Asn .	Ala	Thr	Glu	Ile 1015	Thr	Gln	Phe	Glu	Phe 1020	Pro	Arg	Gln	Lys	3072
tca t Ser S 1025	er 1	Jys .	Arg	Leu	Cys 1030	Ile	Ala	Aap	Phe	Phe 1039	Ala	Pro	Lys	Asp	Ser 1040	3120
gga a Gly I	re 1	.1e /	Asp 1	045	Pne	Pro	Met	Gln	Ala 1050	Val	Thr	Val	Gly :	Glu 1055	Ile	3168
gct ac Ala Ti	nr G	10	Phe 060	Ala	Gln	Lys	Leu	Phe 065	Ala	Asn	Asn	Gln 1	Tyr 1070	Thr	Asp	3216
tat cu Tyr Le	10	75	ne .	H1S	GIA	Leu 1	Ala 080	Val	Gln	Val	Ala 1	Glu 085	Ala	Leu	Ala	3264
gag to Glu Tr 109	р т. 90	nr E	ils /	Ala .	Arg 1	11e 095	Arg	Arg	Glu	Leu 1	Gly 100	Phe	Gly	Ala	Glu	3312
gaa co Glu Pr 1105	O A	sp A	sn .	: 11e	Arg 1	Asp	Ile	Leu	Ala	Gln 1115	Arg '	Tyr	Gln	Gly	Ser 1120	3360
cgg ta Arg Ty	T S	er P	ne (25	(yr)	Pro 1	Ala	Сув 1	Pro 130	Asn	Ile	Gln .	Asp 1	Gln 135	Phe	3408
aag ca Lys Gl	n Le	2U A	sp 1 40	eu I	eu (3lu :	Thr 1	Ser . 145	Arg	Ile i	Asn I	Leu '	Tyr 150	Met	Asp	3456
gaa ag Glu Se	115	55	ın L	eu 1	yr i	ro (31u 6 160	Gln :	Ser '	Thr ?	Thr 1	gcg a Ala : 165	att . Ile	att Ile	act Thr	3504
tat ca Tyr Hi 117	e Pr	a gr	tag al A	ct a la L	ys T	ac t yr I 75	tc a	cc (gcg (Ala	taa						3537

<210> 4 <211> 1178 <212> PRT

<213> Anabaena sp.

<400> 4

Met Thr His Pro Phe Leu Lys Arg Leu His Ser Pro Glu Leu Pro Val 1 5 10 15

Ile Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Thr Gln Asn Leu 20 25 30

Thr Ala Glu Asp Phe Gly Gly Val Gln Tyr Glu Gly Cys Asn Glu Tyr
35 40 45

Leu Val His Thr Lys Pro Glu Ala Val Ala Lys Val His Arg Asp Phe 50 55 60

Leu Ala Val Gly Ala Asp Val Ile Glu Thr Asp Thr Phe Gly Ala Thr
65 70 75 80

Ser Ile Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Thr Tyr Tyr Leu 85 90 95

Asn Lys Lys Ala Ala Glu Leu Ala Lys Ser Val Ala Ala Glu Phe Ser 100 105 110

Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr 115 120 125

Lys Leu Pro Thr Leu Gly His Ile Asp Phe Asp Thr Leu Lys Thr Cys
130 140

Phe Ala Glu Gln Ala Glu Ala Leu Leu Asp Gly Gly Val Asp Leu Leu 145 155 160

Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn 165 170 175

Gly Ile Glu Glu Val Phe Gly Lys Arg Gly Glu Arg Ile Pro Leu Met 180 185 190

Val Ser Val Thr Met Glu Ser Met Gly Thr Met Leu Val Gly Ser Glu 195 200 205

Ile Asn Ala Val Leu Thr Ile Leu Glu Pro Phe Pro Ile Asp Ile Leu 210 215 220

Gly Leu Asn Cys Ala Thr Gly Pro Asp Leu Met Lys Pro His Ile Lys 225 230 235 240

Tyr Leu Ala Glu His Ser Pro Phe Val Val Ser Cys Ile Pro Asn Ala 245 250 255

Gly Leu Pro Glu Asn Val Gly Gly Gln Ala His Tyr Arg Leu Thr Pro 260 265 270

Met Glu Leu Arg Met Ala Leu Met His Phe Val Glu Asp Leu Gly Val 275 280 285

- Gln Val Ile Gly Gly Cys Cys Gly Thr Arg Pro Glu His Ile Gln Gln 290 295 300
- Leu Ala Glu Ile Ala Lys Asp Leu Lys Pro Lys Val Arg Gln Pro Ser 305 310 315 320
- Leu Glu Pro Ala Ala Ala Ser Ile Tyr Ser Thr Gln Pro Tyr Glu Gln 325 330 335
- Asp Asn Ser Phe Leu Ile Val Gly Glu Arg Leu Asn Ala Ser Gly Ser 340 350
- Lys Lys Cys Arg Asp Leu Leu Asn Ala Glu Asp Trp Asp Gly Leu Val 355 360 365
- Ser Met Ala Arg Ser Gln Val Lys Glu Gly Ala His Ile Leu Asp Val 370 380
- Asn Val Asp Tyr Val Gly Arg Asp Gly Val Arg Asp Met His Glu Leu 385 390 395 400
- Val Ser Arg Ile Val Asn Asn Val Thr Leu Pro Leu Met Leu Asp Ser 405 410 415
- Thr Glu Trp Glu Lys Met Glu Ala Gly Leu Lys Val Ala Gly Gly Lys 420 425 430
- Cys Leu Leu Asn Ser Thr Asn Tyr Glu Asp Gly Glu Pro Arg Phe Leu 435 440 445
- Lys Val Leu Glu Leu Ala Lys Lys Tyr Gly Ala Gly Val Val Ile Gly
 450 460
- Thr Ile Asp Glu Glu Gly Met Ala Arg Thr Ala Glu Lys Lys Phe Gln 465 470 475 480
- Ile Ala Gln Arg Ala Tyr Arg Gln Ser Val Glu Tyr Gly Ile Pro Pro 495 495
- Thr Glu Ile Phe Phe Asp Thr Leu Ala Leu Pro Ile Ser Thr Gly Ile 500 505 510
- Glu Glu Asp Arg Glu Asn Gly Lys Ala Thr Ile Glu Ser Ile Ser Arg
 515 520 525
- Ile Arg Lys Glu Leu Pro Gly Cys His Val Ile Leu Gly Val Ser Asn 530 535 540
- Ile Ser Phe Gly Leu Asn Ser Ala Ser Arg Met Val Leu Asn Ser Val 545 550 555 560
- Phe Leu His Glu Ala Met Thr Ala Gly Met Asp Ala Ala Ile Val Ser 565 570 575
- Ala Ser Lys Ile Leu Pro Leu Ser Lys Ile Glu Glu Arg His Gln Glu 580 585 590
- Val Cys Arg Gln Leu Ile Tyr Asp Gln Arg Lys Phe Glu Gly Asp Ile 595 600 605
- Cys Ile Tyr Asp Pro Leu Thr Glu Leu Thr Lys Leu Phe Glu Gly Val

WO 03/087386 17 620 Thr Thr Lys Arg Asn Lys Gly Val Asp Glu Ser Leu Pro Ile Glu Glu 630 635 Arg Leu Lys Arg His Ile Ile Asp Gly Glu Arg Ile Gly Leu Glu Ala Gln Leu Thr Lys Ala Leu Glu Gln Tyr Pro Pro Leu Glu Ile Ile Asn 660 665 Thr Phe Leu Leu Asp Gly Met Lys Val Val Gly Glu Leu Phe Gly Ser 680 Gly Gln Met Gln Leu Pro Phe Val Leu Gln Ser Ala Glu Thr Met Lys Ala Ala Val Ala Tyr Leu Glu Pro Phe Met Glu Lys Ser Glu Ser Gly Asn Asn Ala Lys Gly Lys Val Ile Ile Ala Thr Val Lys Gly Asp Val 725 His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly 745 Tyr Lys Val Ile Asn Leu Gly Ile Lys Gln Pro Val Glu Asn Ile Ile Glu Ala Tyr Asn Gln His Lys Ala Asp Cys Ile Ala Met Ser Gly Leu Leu Val Lys Ser Thr Ala Phe Met Lys Glu Asn Leu Glu Val Phe Asn 790 810 Pro Lys Phe Val His Lys Asp Cys Gln Asn Thr Tyr Lys Gly Lys Val

Glu Lys Gly Ile Asn Val Pro Val Ile Leu Gly Gly Ala Ala Leu Thr

825

Ile Tyr Gly Lys Asp Ala Phe Ser Asp Leu His Phe Met Asp Lys Leu

Met Pro Ala Lys Ala Thr Gly Lys Trp Asp Asn Ser Leu Gly Phe Leu 855

Asp Glu Val Glu Thr Glu Glu Thr Glu Pro Thr Asn His Lys Ser Pro 870

Ile Pro Ser Pro Gln Ser Pro Val Pro Ser Pro Gln Ser Pro Val Pro 890

Ile Asp Thr Arg Arg Ser Glu Ala Val Ala Ile Asp Ile Pro Arg Pro

Thr Pro Pro Phe Trp Gly Thr Gln Leu Leu Gln Pro Ser Asp Ile Ser

Leu Glu Glu Ile Phe Trp His Met Asp Leu Gln Ala Leu Ile Ala Gly 930 940

WO 03/087386 PCT/EP03/04010

4	^
1	×

Gln Trp Gln Phe Arg Lys Pro Lys Glu Gln Ser Lys Glu Glu Tyr Gln 945 950 955 960

Ala Phe Leu Asn Glu Lys Val Tyr Pro Val Leu Glu Thr Trp Lys Gln
965 970 975

Arg Ile Ile Ala Glu Asn Leu Leu His Pro Gln Val Ile Tyr Gly Tyr 980 985 990

Phe Pro Cys Gln Ser Glu Gly Asn Thr Leu Tyr Val Tyr Glu Thr Asn 995 1000 1005

Ser Pro Asn Ala Thr Glu Ile Thr Gln Phe Glu Phe Pro Arg Gln Lys

Ser Ser Lys Arg Leu Cys Ile Ala Asp Phe Phe Ala Pro Lys Asp Ser 1025 1030 1035 1040

Gly Ile Ile Asp Val Phe Pro Met Gln Ala Val Thr Val Gly Glu Ile 1045 1050 1055

Ala Thr Glu Phe Ala Gln Lys Leu Phe Ala Asn Asn Gln Tyr Thr Asp 1060 1065 1070

Tyr Leu Tyr Phe His Gly Leu Ala Val Gln Val Ala Glu Ala Leu Ala 1075 1080 1085

Glu Trp Thr His Ala Arg Ile Arg Arg Glu Leu Gly Phe Gly Ala Glu 1090 1095 1100

Glu Pro Asp Asn Ile Arg Asp Ile Leu Ala Gln Arg Tyr Gln Gly Ser 1105 1110 1115 1120

Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn Ile Gln Asp Gln Phe 1125 1130 1135

Lys Gln Leu Asp Leu Leu Glu Thr Ser Arg Ile Asn Leu Tyr Met Asp 1140 1145 1150

Glu Ser Glu Gln Leu Tyr Pro Glu Gln Ser Thr Thr Ala Ile Ile Thr 1155 1160 1165

Tyr His Pro Val Ala Lys Tyr Phe Thr Ala 1170 1175

<210> 5

<211> 3588

<212> DNA

<213> Synechocystis sp.

<220>

<221> CDS

<222> (1)..(3585)

<223> RCY35965

<400> 5

atg aaa agt gct ttt tta gac cgt atc cac agt ccc gat cgc ccg gta

Met Lys Ser Ala Phe Leu Asp Arg Ile His Ser Pro Asp Arg Pro Val

1 5 10 15

tta gtc ttt gac ggg gct atg ggt aca aac ctg cag gta cag aac cta 96

										19				•		
Lev	l Val	. Phe	20	Gly	Ala	Met	Gly	Thr 25			Gln	Val	Gln 30		Leu	
acg Thr	dca Ala	gcg Ala 35	Asp	ttt Phe	ggt Gly	999 Gly	gcg Ala 40	Glu	tac Tyr	gaa Glu	ggt Gly	tgc Cys 45	. aat Asn	gaa Glu	tat Tyr	144
tta Leu	gtc Val 50	His	acc Thr	aag Lys	cca Pro	gag Glu 55	gcc Ala	gtg Val	gct Ala	acg Thr	gtg Val 60	cat His	cgt Arg	gct Ala	ttt Phe	192
tac Tyr 65	Glu	gcg	ggg Gly	gcc	gat Asp 70	gtc Val	gtg Val	gaa Glu	acg Thr	gat Asp 75	act Thr	ttt Phe	ggg Gly	gga Gly	acg Thr 80	240
Pro	Leu	Val	Leu	Ala 85	gag Glu	Tyr	Asp	Leu	Ala 90	Asp	Gln	Ser	Tyr	Tyr 95	Leu	288
Asn	Lys	Ala	Ala 100	Ala	gag Glu	Leu	Ala	Lys 105	Ala	Val	Ala	Ala	Glu 110	Phe	Ser	336
Thr	Pro	Glu 115	Lys	Pro	cga Arg	Phe	Val 120	Ala	Gly	Ser	Met	Gly 125	Pro	Gly	Thr	384
Lys	Leu 130	Pro	Thr	Leu	ggt Gly	His 135	Val	qaA	Tyr	Asp	Ser 140	Leu	Lys	Asp	Ala	432
145	Val	Val	Gln	Val	cgg Arg 150	Gly	Leu	Tyr	Авр	Gly 155	Gly	Val	Asp	Leu	Leu 160	480
Leu	Val	Glu	Thr	Сув 165	cag Gln	Авр	Val	Leu	Gln 170	Ile	Lys	Ala	Ala	Leu 175	Asn	528
Ala	Ile	Glu	Gln 180	Val	ttt Phe	Ala	Glu	Lys 185	Gly	Asp	Arg	Leu	Pro 190	Leu	Met	576
gtg Val	tca Ser	gta Val 195	acc Thr	atg Met	gaa Glu	acc Thr	atg Met 200	G1y 999	acc Thr	atg Met	ctg Leu	gtg Val 205	ggt Gly	acg Thr	gag Glu	624
atg Met	gcg Ala 210	gcg Ala	gcc Ala	ctg Leu	gcc Ala	att Ile 215	ttg Leu	gag Glu	ccc Pro	tat Tyr	ccc Pro 220	atc Ile	gat Asp	att Ile	ttg Leu	672
999 Gly 225	cta Leu	aac Asn	tgc Cys	gcc Ala	acc Thr 230	999 Gly	cca Pro	gat Asp	ttg Leu	atg Met 235	aag Lys	gaa Glu	cac His	gtt Val	aaa Lys 240	720
tat Tyr	ctt Leu	tcc Ser	Glu	cat His 245	tcc Ser	ccc Pro	ttt Phe	Val	gtg Val 250	tcc Ser	tgt Cys	att Ile	ccc Pro	aat Asn 255	gct Ala	768
ggt Gly	ttg Leu	Pro	gaa Glu 260	aac Asn	gtt Val	ggc Gly	ggt Gly	caa Gln 265	gct Ala	ttt Phe	tat Tyr	cgc Arg	ctc Leu 270	acc Thr	ccg Pro	816

a M	tg et	gaa Glu	Let 275	ı Glı	a ato 1 Met	tco Ser	ctg Leu	atg Met 280	His	ttc Phe	ato : Ile	gaa Glu	gac Asp 285	Leu	gga Gly	gta Val	864
C G	ag ln	gta Val 290	Πle	ggt Gly	ggt Gly	tgt Cys	tgt Cys 295	Gly	act Thr	aga Arg	ccc Pro	gat Asp 300	His	atc	aag Lys	gcc Ala	912
L	tg eu 05	gcg Ala	gat Asp	att Ile	gcc Ala	aag Lys 310	gat Asp	ctc	cag Gln	ccc Pro	aaa Lys 315	caa Gln	cgc Arg	caa Gln	cct Pro	cac His 320	960
t: T	ac yr	gaa Glu	Pro	agc Ser	gcc Ala 325	Ala	tcc Ser	att Ile	tat Tyr	tcc Ser 330	Thr	caa Gln	acc Thr	tac Tyr	gcc Ala 335	caa Gln	1008
G:	aa lu	aat Asn	tct Ser	Phe 340	Leu	atc Ile	att Ile	ggc Gly	gaa Glu 345	cgg Arg	ctc Leu	aat Asn	gcc Ala	agt Ser 350	ggc Gly	tcg Ser	1056
Ly	/8	Lys	Cys 355	Arg	Asp	Leu	ctc Leu	Asn 360	Ala	Glu	Asp	Trp	Asp 365	Ser	Leu	Val	1104
S€	er	Leu 370	Ala	Lys	Ser	Gln	gtc Val 375	Lys	Glu	Gly	Ala	Gln 380	Ile	Leu	Asp	Val	1152
As 38	5 5	Val	Asp	Tyr	Val	Gly 390	cga Arg	Asp	Gly	Val	Arg 395	Asp	Met.	Lys	Glu	Leu 400	1200
Al	a	Ser	Arg	Leu	Val 405	Asn	aat Asn	Val	Thr	Leu 410	Pro	Leu	Met	Leu	Asp 415	Ser	1248
Th	r	Glu	Trp	Gln 420	Lys	Met	gag Glu	Ala	Gly 425	Leu	Lys	Val	Ala	Gly 430	Gly	Lys	1296
Су	s :	Ile	Leu 435	Asn	Ser	Thr	aac Asn	Tyr 440	Glu	qaA	Gly	Glu	Glu 445	Arg	Phe	Tyr	1344
Ly	8 \	Val 450	Leu	Glu	Ile	Ala	aaa Lys 455	Glu	Tyr	Gly	Ala	Gly 460	Ile	Val	Ile	Gly	1392
Th 46	r] 5	lle	Asp	Glu	Asp	Gly 470	atg Met	Gly .	Arg	Thr	Ala 475	Asp	Lys	Lys	Phe	Glu 480	1440
11	e A	\la	Lys	Arg	Ala 485	Tyr	gaa Glu	Ala .	Ala	Ile 490	Ala	Phe	Gly	Ile	Pro 495	Ala	1488
Th	rG	3lu	Ile	Phe 500	Phe	qaA	cct Pro	Leu :	Ala 505	Leu	Pro	Ile	Ser	Thr 510	Gly	Ile	1536
gaa	a g	jaa 🤈	gac	agg	gag	aac	ggt	aaa 🤉	gcc	acc	gtg	gat	gct	atc	cgc	aga	1584

Glu	Glu	Asp 515	Arg	, Glu	Asn	Gly	Lys 520			21 Val	Asp	Ala 525		Arg	Arg	
att Ile	cgc Arg	Gln	gaa Glu	ttg Leu	Pro	gat Asp 535	tgt Cys	cat His	att Ile	ttg Leu	ttg Leu 540	99 9 Gly	gtt Val	tct Ser	aac Asn	1632
gtt Val 545	Ser	ttt Phe	ggc	ttg Leu	aat Asn 550	Pro	gcc Ala	gct Ala	cgc Arg	cag Gln 555	gta Val	ctc Leu	aat Asn	tcc Ser	atc Ile 560	1680
ttt Phe	ctc	cac His	gaa Glu	tgt Cys 565	atg Met	cag Gln	gtg Val	Gly	atg Met 570	gat Asp	gcg Ala	gcc Ala	att Ile	gtc Val 575	agt Ser	1728
gcc Ala	aat Asn	aag Lys	att Ile 580	tta Leu	ccc Pro	ctg Leu	gca Ala	aaa Lys 585	att Ile	gac Asp	cca Pro	gaa Glu	caa Gln 590	caa Gln	caa Gln	1776
gtc Val	tgt Cys	cta Leu 595	gat Asp	tta Leu	atc Ile	tat Tyr	gac Asp 600	cgc Arg	cgg Arg	gaa Glu	ttt Phe	gaa Glu 605	gga Gly	gag Glu	cgc Arg	1824
tgt Cys	aca Thr 610	tat Tyr	gac Asp	ccg Pro	tta Leu	acc Thr 615	aaa Lys	ctc Leu	acc Thr	act Thr	tta Leu 620	ttt Phe	gaa Glu	ggt Gly	aaa Lys	1872
acc Thr 625	acc Thr	aaa Lys	cgg Arg	gat Asp	aaa Lys 630	tcc Ser	ggt Gly	gat Asp	gcc Ala	aat Asn 635	tta Leu	ccg Pro	gtg Val	gaa Glu	gaa Glu 640	1920
aga Arg	tta Leu	aaa Lys	cgc Arg	cac His 645	atc Ile	att Ile	gat Asp	61y 999	gaa Glu 650	aga Arg	ttg Leu	ggc Gly	tta Leu	gaa Glu 655	gag Glu	1968
gcc Ala	ctc Leu	aat Asn	gaa Glu 660	gct Ala	tta Leu	aaa Lys	ctt Leu	tac Tyr 665	gct Ala	ccc Pro	tta Leu	gat Asp	atc Ile 670	att Ile	aac Asn	2016
atc Ile	tat Tyr	ttg Leu 675	ttg Leu	gat Asp	ggc Gly	atg Met	aaa Lys 680	gtg Val	gtg Val	G1y 999	gaa Glu	cta Leu 685	ttt Phe	ggt Gly	tcc Ser	2064
999 Gly	caa Gln 690	atg Met	cag Gln	ttg Leu	ccc Pro	ttt Phe 695	gtg Val	ttg Leu	cag Gln	tcg Ser	gcc Ala 700	caa Gln	acc Thr	atg Met	aaa Lys	2112
gcg Ala 705	gcg Ala	gtg Val	gct Ala	ttt Phe	tta Leu 710	gaa Glu	ccc Pro	cat His	atg Met	gat Asp 715	aag Lys	gat Asp	gat Asp	tcc Ser	gcc Ala 720	2160
gac Asp	aat Asn	gct Ala	aag Lys	ggt Gly 725	act Thr	ttt Phe	tta Leu	att Ile	gcc Ala 730	act Thr	gtt Val	aag Lys	GJY 999	gat Asp 735	gtc Val	2208
cat His	gat Asp	att Ile	ggc Gly 740	aaa Lys	aac Asn	tta Leu	gtg Val	gat Asp 745	att Ile	atc Ile	ctt Leu	tcc Ser	aac Asn 750	aat Asn	ggc Gly	2256
tat Tyr	cga Arg	gtg Val 755	gtc Val	aac Asn	cta Leu	Gly	att Ile 760	aaa Lys	cag Gln	cca Pro	Val	gaa Glu 765	aat Asn	att Ile	atc Ile	2304

gaa Glu	gcc Ala 770	Tyr	aaa Lys	aaa Lys	cac His	agg Arg 775	Pro	gat Asp	tgc Cys	att Ile	gcc Ala 780	Met	agt Ser	ggt Gly	ttg Leu	2352
ttg Leu 785	Val	aaa Lys	tca Ser	act Thr	gct Ala 790	Phe	atg Met	aag Lys	gaa Glu	aat Asn 795	Leu	gaa Glu	gtt Val	ttc Phe	aac Asn 800	2400
caa Gln	gag Glu	ggc	att	act Thr 805	gtt Val	ccc Pro	gtc Val	att Ile	ctt Leu 810	ggt Gly	ggt	gct Ala	gct Ala	tta Leu 815	acg Thr	2448
cct Pro	aaa Lys	ttt Phe	gtt Val 820	cac His	cag Gln	gac Asp	tgc Cys	caa Gln 825	aat Asn	acc Thr	tac Tyr	aaa Lys	ggc Gly 830	caa Gln	gta Val	2496
Ile	Tyr	Gly B35	Lys	Asp	Ala	Phe	Ala 840	Asp	Leu	His	Phe	Met 845	Asp	aag Lys	Leu	2544
Met	Pro 850	Ala	Lys	Asn	Ser	His 855	Asn	Trp	Asp	Asp	Phe 860	Gln	Gly	ttt Phe	Leu	2592
Gly 8 6 5	Glu	Tyr	Ala	Thr	Glu 870	Asn	Gly	His	Asn	Val 875	Thr	Thr	Asp	gat Asp	Gly 880	2640
Ala	Lys	Thr	Asn	Phe 885	Gly	Ile	Glu	Glu	Glu 890	Lys	Leu	Ile	Asp	gct Ala 895	Ser	2688
Glu	Gln	Ser	Arg 900	Glu	Pro	Glu	Val	Ile 905	Asp	Thr	Val	Arg	Ser 910	gaa Glu	Ala	2736
Val	Asp	Pro 915	Asp	Leu	Glu	Arg	Pro 920	Val	Pro	Pro	Phe	Trp 925	Gly	act Thr	Lys	2784
Ile	Leu 930	Gln	Ser	Ser	Asp	11e 935	Ser	Leu	Asp	Glu	Val 940	Phe	Pro	tta Leu	Leu	2832
Asp 945	Leu	Gln	Ala	Leu	Phe 950	Val	Gly	Gln	Trp	Gln 955	Phe	Arg	Lys	cct Pro	Arg 960	2880
Glu	Gln	Ser	Arg	Glu 965	Glu	Tyr	Glu	Gln	Phe 970	Leu	Ala	Glu	Lys	gtt Val 975	His	2928
Pro	Ile	Leu	Ala 980	Glu	Trp	ГÀв	Gly	Lys 985	Val	Met	Ala	Glu	Asn 990	tta Leu	Leu	2976
His	Pro	Thr 995	Val	Val	Tyr	Gly 1	Tyr 000	Phe	Pro	Сув	Gln 1	Ser .005	Gln	ggc Gly	Asn	3024
acc	ttg	tta	att	tat	gac	cca	gaa	ttg	gtc	agc	caa	aat	aat	ggc	caa	3072

										Z J						
	Leu 010	Leu	Ile	Tyr	. yeb	Pro 1015		Leu	Val		Gln 1020		Asn	Gly	Gln	•
att Ile 1025	Pro	cca Pro	gac Asp	gca Ala	acg Thr 103	Ala	atc Ile	gcc Ala	aaa Lys	ttt Phe 103	Glu	ttt Phe	Pro	.cgg Arg	caa Gln 1040	3120
aaa Lys	tca Ser	ggg ggg	Arg	cgg Arg 1045	Leu	tgt Cys	att Ile	Ala	gac Asp 1050	Phe	ttt Phe	gct Ala	Ser	aaa Lys 1055	gaa Glu	3168
tcg (61 y 999	Ile	act Thr 1060	gat Asp	gtt Val	ttt Phe	Pro	ttg Leu 1065	caa Gln	gcg Ala	gtt Val	Thr	gtg Val 1070	61y 999	gaa Glu	3216
atc (Ala	acg Thr 1075	gaa Glu	tat Tyr	gca Ala	Arg	aaa Lys 1080	ctt Leu	ttt Phe	gct Ala	Gly	gat Asp 1085	aat Asn	tac Tyr	acc Thr	3264
gat (tac Tyr 090	ctc Leu	tac Tyr	ttc Phe	His	ggc Gly 1095	atg Met	gcg Ala	gtg Val	Gln	atg Met 1100	gcg Ala	gaa Glu	gct Ala	tta Leu	3312
gcg g Ala (1105	gag Glu	tgg Trp	act Thr	cac His	caa Gln 1110	Arg	ata Ile	cgt Arg	cag Gln	gaa Glu 1115	Leu	ggc Gly	ttt Phe	ggc Gly	cat His 1120	3360
tta g Leu <i>l</i>	gat Asp	cca Pro	qaA	aac Asn 125	atc Ile	cgt Arg	gat Asp	Leu	ctc Leu 1130	cag Gln	caa Gln	cgt Arg	Tyr	caa Gln 1135	ggt Gly	3408
tcc o	age Arg	Tyr	agt Ser 140	ttt Phe	ggt Gly	tat Tyr	Pro	gct Ala 1145	tgt Cys	ccc Pro	aac Asn	Met	cag Gln 1150	gat Asp	caa Gln	3456
tac a Tyr 1	hr	caa Gln 155	tta Leu	gaa Glu	ttg Leu	Leu	caa Gln 160	acc Thr	gaa Glu	cga Arg	Ile	ggc Gly 1165	ttg Leu	tat Tyr	atg Met	3504
gat g Asp G	jaa 31u 170	agt Ser	gaa Glu	cag Gln	Val	tat Tyr 175	cca Pro	gaa Glu	caa Gln	Ser	acc Thr 180	acg Thr	gcg Ala	att Ile	att Ile	3552
tcc t Ser T 1185	at Yr	cat His	cct Pro	gcg Ala	gct Ala 1190	Lys	tat Tyr	ttc Phe	agc Ser	gct Ala 1195						3588
<210><211><211><212><213>	11 PR	T	ocys	tis	вр.											
<400> Met L 1		Ser .	Ala	Phe 5	Leu .	qeA	Arg	Ile	His 10	Ser	Pro	Asp	Arg	Pro 15	Val	

Leu Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Val Gln Asn Leu 20 25 30

Thr Ala Ala Asp Phe Gly Gly Ala Glu Tyr Glu Gly Cys Asn Glu Tyr

Leu Val His Thr Lys Pro Glu Ala Val Ala Thr Val His Arg Ala Phe
50 60

Tyr Glu Ala Gly Ala Asp Val Val Glu Thr Asp Thr Phe Gly Gly Thr
65 70 75 80

Pro Leu Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Ser Tyr Tyr Leu 85 90

Asn Lys Ala Ala Ala Glu Leu Ala Lys Ala Val Ala Ala Glu Phe Ser 100 105 110

Thr Pro Glu Lys Pro Arg Phe Val Ala Gly Ser Met Gly Pro Gly Thr

Lys Leu Pro Thr Leu Gly His Val Asp Tyr Asp Ser Leu Lys Asp Ala 130 135 140

Tyr Val Val Gln Val Arg Gly Leu Tyr Asp Gly Gly Val Asp Leu Leu 145 150 155 160

Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn 165 170 175

Ala Ile Glu Gln Val Phe Ala Glu Lys Gly Asp Arg Leu Pro Leu Met 180 185 190

Val Ser Val Thr Met Glu Thr Met Gly Thr Met Leu Val Gly Thr Glu 195 200 205

Met Ala Ala Leu Ala Ile Leu Glu Pro Tyr Pro Ile Asp Ile Leu 210 215 220

Gly Leu Asn Cys Ala Thr Gly Pro Asp Leu Met Lys Glu His Val Lys 225 230 235 240

Tyr Leu Ser Glu His Ser Pro Phe Val Val Ser Cys Ile Pro Asn Ala 245 250 255

Gly Leu Pro Glu Asn Val Gly Gly Gln Ala Phe Tyr Arg Leu Thr Pro 260 265 270

Met Glu Leu Gln Met Ser Leu Met His Phe Ile Glu Asp Leu Gly Val 275 280 285

Gln Val Ile Gly Gly Cys Cys Gly Thr Arg Pro Asp His Ile Lys Ala 290 295 300

Leu Ala Asp Ile Ala Lys Asp Leu Gln Pro Lys Gln Arg Gln Pro His 305 310 315 320

Tyr Glu Pro Ser Ala Ala Ser Ile Tyr Ser Thr Gln Thr Tyr Ala Gln 325 330 335

Glu Asn Ser Phe Leu Ile Ile Gly Glu Arg Leu Asn Ala Ser Gly Ser 340 345 350

Lys Lys Cys Arg Asp Leu Leu Asn Ala Glu Asp Trp Asp Ser Leu Val 355 360 365

- Ser Leu Ala Lys Ser Gln Val Lys Glu Gly Ala Gln Ile Leu Asp Val 370 380
- Asn Val Asp Tyr Val Gly Arg Asp Gly Val Arg Asp Met Lys Glu Leu 385 390 395 400
- Ala Ser Arg Leu Val Asn Asn Val Thr Leu Pro Leu Met Leu Asp Ser
 405 410 415
- Thr Glu Trp Gln Lys Met Glu Ala Gly Leu Lys Val Ala Gly Gly Lys
 420 425 430
- Cys Ile Leu Asn Ser Thr Asn Tyr Glu Asp Gly Glu Glu Arg Phe Tyr 435 440 445
- Lys Val Leu Glu Ile Ala Lys Glu Tyr Gly Ala Gly Ile Val Ile Gly 450 455 460
- Thr Ile Asp Glu Asp Gly Met Gly Arg Thr Ala Asp Lys Lys Phe Glu 465 470 475 480
- Ile Ala Lys Arg Ala Tyr Glu Ala Ala Ile Ala Phe Gly Ile Pro Ala 485 490 495
- Thr Glu Ile Phe Phe Asp Pro Leu Ala Leu Pro Ile Ser Thr Gly Ile 500 505 510
- Glu Glu Asp Arg Glu Asn Gly Lys Ala Thr Val Asp Ala Ile Arg Arg 515 520 525
- Ile Arg Gln Glu Leu Pro Asp Cys His Ile Leu Leu Gly Val Ser Asn 530 535 540
- Val Ser Phe Gly Leu Asn Pro Ala Ala Arg Gln Val Leu Asn Ser Ile 545 550 555 560
- Phe Leu His Glu Cys Met Gln Val Gly Met Asp Ala Ala Ile Val Ser 565 570 575
- Ala Asn Lys Ile Leu Pro Leu Ala Lys Ile Asp Pro Glu Gln Gln 585 590
- Val Cys Leu Asp Leu Ile Tyr Asp Arg Arg Glu Phe Glu Gly Glu Arg 595 600 605
- Cys Thr Tyr Asp Pro Leu Thr Lys Leu Thr Thr Leu Phe Glu Gly Lys 610 620
- Thr Thr Lys Arg Asp Lys Ser Gly Asp Ala Asn Leu Pro Val Glu Glu 625 635 640
- Arg Leu Lys Arg His Ile Ile Asp Gly Glu Arg Leu Gly Leu Glu Glu 645 655
- Ala Leu Asn Glu Ala Leu Lys Leu Tyr Ala Pro Leu Asp Ile Ile Asn 660 665 670
- Ile Tyr Leu Leu Asp Gly Met Lys Val Val Gly Glu Leu Phe Gly Ser 675 680 685
- Gly Gln Met Gln Leu Pro Phe Val Leu Gln Ser Ala Gln Thr Met Lys 690 695 700

- Ala Ala Val Ala Phe Leu Glu Pro His Met Asp Lys Asp Asp Ser Ala 705 710 715 720
- Asp Asn Ala Lys Gly Thr Phe Leu Ile Ala Thr Val Lys Gly Asp Val 725 730 735
- His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly 740 745 750
- Tyr Arg Val Val Asn Leu Gly Ile Lys Gln Pro Val Glu Asn Ile Ile 755 760 765
- Glu Ala Tyr Lys Lys His Arg Pro Asp Cys Ile Ala Met Ser Gly Leu 770 780
- Leu Val Lys Ser Thr Ala Phe Met Lys Glu Asn Leu Glu Val Phe Asn 785 790 795 800
- Gln Glu Gly Ile Thr Val Pro Val Ile Leu Gly Gly Ala Ala Leu Thr 805 810 815
- Pro Lys Phe Val His Gln Asp Cys Gln Asn Thr Tyr Lys Gly Gln Val 820 825 830
- Ile Tyr Gly Lys Asp Ala Phe Ala Asp Leu His Phe Met Asp Lys Leu 835 840 845
- Met Pro Ala Lys Asn Ser His Asn Trp Asp Asp Phe Gln Gly Phe Leu 850 855 860
- Gly Glu Tyr Ala Thr Glu Asn Gly His Asn Val Thr Thr Asp Asp Gly 865 870 875 880
- Ala Lys Thr Asn Phe Gly Ile Glu Glu Glu Lys Leu Ile Asp Ala Ser 885 890 895
- Glu Gln Ser Arg Glu Pro Glu Val Ile Asp Thr Val Arg Ser Glu Ala 900 905 910
- Val Asp Pro Asp Leu Glu Arg Pro Val Pro Pro Phe Trp Gly Thr Lys 915 920 925
- Ile Leu Gln Ser Ser Asp Ile Ser Leu Asp Glu Val Phe Pro Leu Leu 930 935 940
- Asp Leu Gln Ala Leu Phe Val Gly Gln Trp Gln Phe Arg Lys Pro Arg 945 950 955 960
- Glu Gln Ser Arg Glu Glu Tyr Glu Gln Phe Leu Ala Glu Lys Val His 965 970 975
- Pro Ile Leu Ala Glu Trp Lys Gly Lys Val Met Ala Glu Asn Leu Leu 980 985 990
- His Pro Thr Val Val Tyr Gly Tyr Phe Pro Cys Gln Ser Gln Gly Asn 995 1000 1005
- Thr Leu Leu Ile Tyr Asp Pro Glu Leu Val Ser Gln Asn Asn Gly Gln 1010 1015 1020
- Ile Pro Pro Asp Ala Thr Ala Ile Ala Lys Phe Glu Phe Pro Arg Gln

1025

1030

1035

1040

Lys Ser Gly Arg Arg Leu Cys Ile Ala Asp Phe Phe Ala Ser Lys Glu 1045 1050 1055

Ser Gly Ile Thr Asp Val Phe Pro Leu Gln Ala Val Thr Val Gly Glu 1060 1065 1070

Ile Ala Thr Glu Tyr Ala Arg Lys Leu Phe Ala Gly Asp Asn Tyr Thr 1075 1080 ... 1085 .

Asp Tyr Leu Tyr Phe His Gly Met Ala Val Gln Met Ala Glu Ala Leu 1090 1095 1100

Ala Glu Trp Thr His Gln Arg Ile Arg Gln Glu Leu Gly Phe Gly His 1105 1110 1115 1120

Leu Asp Pro Asp Asn Ile Arg Asp Leu Leu Gln Gln Arg Tyr Gln Gly
1125 1130 1135

Ser Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn Met Gln Asp Gln 1140 1145 1150

Tyr Thr Gln Leu Glu Leu Leu Gln Thr Glu Arg Ile Gly Leu Tyr Met 1155 1160 1165

Asp Glu Ser Glu Gln Val Tyr Pro Glu Gln Ser Thr Thr Ala Ile Ile 1170 1175 1180

Ser Tyr His Pro Ala Ala Lys Tyr Phe Ser Ala 1185 1190 1195

<210> 7

<211> 3561

<212> DNA

<213> Prochlorococcus marinus

<220>

<221> CDS

<222> (1) .. (3558)

<223> RCK00830

<400> 7

atg gtt tca ttt aga aat tat tta aat aga gat gat aaa cca att att 48
Met Val Ser Phe Arg Asn Tyr Leu Asn Arg Asp Asp Lys Pro Ile Ile
1 5 10 15

att ttc gat ggt ggg aca ggt act tct ttt caa aat tta aat tta tca 96
Ile Phe Asp Gly Gly Thr Gly Thr Ser Phe Gln Asn Leu Asn Leu Ser
20 25 30

tca cat gat ttt ggt gga gat gat tta gag ggt tgc aat gaa aac tta 144 Ser His Asp Phe Gly Gly Asp Asp Leu Glu Gly Cys Asn Glu Asn Leu 35

gtt cta tcc tct cct aat act gtt gaa caa gta cat aat tca ttt ctt 192 Val Leu Ser Ser Pro Asn Thr Val Glu Gln Val His Asn Ser Phe Leu 50 60

gaa gca ggt tgt cat gta att gaa acc aat aca ttt ggt gct tca tct 240 Glu Ala Gly Cys His Val Ile Glu Thr Asn Thr Phe Gly Ala Ser Ser

										28				•		
65	i				70)				75					80	
		tta	gac	gaa			att	tct	aat			tat	qaa	atc		288
Ile	· Val	Leu	Asp	Glu 85	Tyr	Ser	Ile	Ser	Asn 90	Lys	Ala	Tyr	Glu	Ile 95	Asn	
aaa Lys	aaa Lys	gca Ala	gct Ala 100	Gln	ata Ile	gct	aaa Lys	aaa Lys 105	tgt Cys	gca Ala	aat Asn	tta Leu	Phe	tca Ser	tct Ser	336
att Ile	aat Asn	act Thr 115	cct Pro	aga Arg	ttt Phe	gtc Val	gct Ala 120	gga Gly	tca Ser	att	613 888	cca Pro 125	act Thr	aca Thr	aaa Lys	384
tta Leu	cca Pro 130	aca Thr	tta Leu	ggt Gly	cat His	att Ile 135	agt Ser	ttt Phe	gat Asp	aag Lys	ctt Leu 140	aaa Lys	gat Asp	tca Ser	tat Tyr	432
gaa Glu 145	gaa Glu	caa Gln	ata Ile	aat Asn	ggt Gly 150	cta Leu	att Ile	gac Asp	gga Gly	ggt Gly 155	att Ile	gac Asp	ctt Leu	cta Leu	ttg Leu 160	480
Ile	Glu	aca Thr	Сув	Gln 165	Asp	Val	Leu	Gln	Ile 170	Lys	Ser	Āla	Leu	Ser 175	Ala	528
Ser	Gln	gaa Glu	Val 180	Ile	Lys	Asn	Arg	Asn 185	Ile	Glu	Leu	Pro	Ile 190	Met	Ile	576
Ser	Ile	act Thr 195	Met	Glu	Thr	Thr	Gly 200	Thr	Met	Leu	Val	Gly 205	Ser	Asp	Ile	624
Ala	Ser 210	gca Ala	Leu	Thr	Ile	Leu 215	Glu	Pro	Tyr	Asn	11e 220	Asp	Ile	Leu	Gly	672
Leu 225	Asn	tgt Cys	Ala	Thr	Gly 230	Pro	Val	Gln	Met	Lys 235	Glu	His	Ile	Lys	Tyr 240	720
Leu	Ala	gaa Glu	Asn	Ser 245	Pro	Phe	Ala	Ile	Ser 250	Сув	Ile	Pro	Asn	Ala 255	Gly	768
Leu	Pro	gaa Glu	Asn 260	Ile	Gly	Gly	Val	Ala 265	His	Tyr	Lys	Leu	Thr 270	Pro	Leu	816
Glu	Leu	aaa Lys 275	Met	Gln	Leu	Met	Asn 280	Phe	Ile	Tyr	Asp	Phe 285	Asn	Val	Gln	864
ctt Leu	att Ile 290	ggc Gly	gga Gly	tgt Cys	Сув	ggt Gly 295	act Thr	act Thr	cct Pro	gaa Glu	cat His 300	atc Ile	aag Lys	cat His	tta Leu	912
tca Ser 305	tca Ser	atc Ile	att Ile	gag Glu	gaa Glu 310	ata Ile	gtt Val	gat Asp	aaa Lys	aaa Lys 315	ata Ile	aat Asn	aaa Lys	aga Arg	ctt Leu 320	960

										29						
Pro	Thr	· Val	l Lys	325	: Asr	Phe	· Val	. Pro	330	Ala	Ala	Ser	Ile	335		1008
gca Ala	gtt Val	cca Pro	tat Tyr 340	Lye	caa Gln	gat Asp	aac Asn	tca Ser 345	Ile	tta Leu	ata Ile	gtt Val	gga Gly 350	Glu	cgt	1056
tta Leu	aat Asn	gct Ala 355	agt Ser	gga Gly	tca Ser	aaa Lys	Lys 360	Val	agg Arg	gaa Glu	tta Leu	cta Leu 365	aat Asņ	gaa Glu	gat Asp	1104
gat Asp	tgg Trp 370	Asp	ggc Gly	ctg Leu	cta Leu	tca Ser 375	att Ile	gct Ala	aaa Lys	caa Gln	cag Gln 380	caa Gln	aaa Lys	gaa Glu	aat Asn	1152
gct Ala 385	His	ata	cta Leu	gat Asp	gtc Val 390	aat Asn	gtt Val	gat Asp	tat Tyr	gta Val 395	gga Gly	aga Arg	gat Asp	gga Gly	gtt Val 400	1200
aaa Lys	gat Aвр	atg Met	aaa Lys	gaa Glu 405	att Ile	acc	tca Ser	aga Arg	tta Leu 410	gtt Val	aca Thr	aat Asn	ata Ile	aat Asn 415	ctt Leu	1248
Pro	Leu	Met	ata Ile 420	Asp	Ser	Thr	Glu	Ala 425	Авр	Lys	Met	Glu	Ser 430	Gly	Leu	1296
Lys	Thr	Val 435	gga Gly	Gly	Гув	Сув	Ile 440	Ile	Asn	Ser	Thr	Asn 445	Tyr	Glu	Asp	1344
Gly	Asp 450	Asp	aga Arg	Phe	Asn	Gln 455	Val	Leu	Arg	Leu	Ala 460	Leu	Asp	Tyr	Gly	1392
Ala 465	Gly	Ile	gta Val	Ile	Gly 470	Thr	Ile	Asp	Glu	Asp 475	Gly	Met	Ala	Arg	Thr 480	1440
Ser	Gln	Lys	aaa Lys	Tyr 485	Asp	Ile	Ala	Lys	Arg 490	Ala	Leu	Ile	Lys	Thr 495	Arg	1488
Ser	Ser	Gly	ctc Leu 500	Ala	Asp	Tyr	Glu	Ile 505	Phe	Phe	Asp	Pro	Leu 510	Āla	Leu	1536
Pro	Ile	Ser 515	act Thr	Gly	Ile	Glu	Glu 520	qaA	Arg	Leu	Asn	Ala 525	ГÀв	Ala	Thr	1584
Ile	Glu 530	Ala	ata Ile	Ser	Lys	11e 535	Arg	Lys	Ser	Phe	Pro 540	двр	Ile	His	Ile	1632
11e 545	Leu	Gly	ata Ile	Ser	Asn 550	Ile	Ser	Phe	Gly	Leu 555	Ser	Pro	Leu	Ser	Arg 560	1680
att Ile	aat Asn	cta Leu	aat Asn	tca Ser	ata Ile	ttt Phe	ctc Leu	gat Asp	gaa Glu	tgt Cys	ata Ile	aag Lys	gca Ala	gga Gly	tta Leu	1728

				565	5				570	30				575	•	
gat Asp	tca Ser	gcg Ala	att Ile 580	: Ile	gca Ala	cca Pro	aat Asn	aaa Lys 585	Ile	ttg Leu	cct Pro	ctt Lev	tca Ser 590	aaa Lys	ata Ile	1776
tct Ser	gcg	gaa Glu 595	Thr	aaa Lys	aaa Lys	tta Leu	tgt Cys 600	Leu	gat Asp	tta Leu	att Ile	tat Tyr 605	Asp	aga Arg	aga Arg	1824
aat Asn	ttc Phe 610	GIU	aat Asn	gaa Glu	ata Ile	tgt Cys 615	Ile	tat	gat	cca Pro	tta Leu 620	Val	gaa Glu	cta Leu	aca Thr	1872
aaa Lys 625	gca Ala	ttc Phe	caa Gln	gat Asp	ata Ile 630	Thr	atc Ile	agt Ser	gac Asp	Phe 635	Lys	aaa Lys	gga Gly	tct Ser	act Thr 640	1920
Ser	Asn	Lys	Asn	645		Leu	Glu	Glu	Б ув	Leu	Lys	Asn	His	Ile 655	Val	1968
Asp	Gly	Glu	660 Lys	Ile	ggt Gly	Leu	Glu	Glu 665	Gln	Leu	Asn	Asn	Ala 670	Leu	Lys	2016
Lys	Tyr	Lys 675	Pro	Leu	gaa Glu	Ile	11e 680	Asn	Thr	Tyr	Leu	Leu 685	Asp	Gly	Met	2064
Lys	Val 690	Val	Gly	Glu	cta Leu	Phe 695	Gly	Ser	Gly	Gln	Met 700	Gln	Leu	Pro	Phe	2112
gta Val 705	Leu	Gln	Ser	Ala	Glu 710	Thr	Met	Lys	Phe	Ala 715	Val	Ser	Val	Leu	Glu 720	2160
cct Pro	His	Met	Glu	Thr 725	Val	Asp	Glu	Lys	11e 730	Ser	Asn	Gly	Lys	Leu 735	Leu	2208
ata (AIA	Tnr	740	Lys	Gly	Asp	Val	His 745	Asp	Ile	Gly	Lув	Asn 750	Leu	Val	2256
gat Asp	Ile	11e 755	Leu	Ser	Asn	Asn	Gly 760	Phe	Asp	Val	Ile	Asn 765	Leu	Gly	Ile	2304
	31n 770	Asp	Val	Ser	Ala	11e 775	Ile	Asp	Ala	Gln	Lys 780	Lys	His	Lys	Ala	2352
gac (Asp (785	Cys	Ile	Ala	Met	Ser 790	Gly	Leu	Leu	Val	Lys 795	Ser	Thr	Ala	Phe	Met 800	2400
aag g Lys A	gat Asp	aat Asn	tta Leu	gaa Glu	gca Ala	ttt Phe	aac Asn	aat Asn	gct Ala	gaa Glu	att Ile	aat Asn	gtt Val	cca Pro	gtt Val	2448

										31							
att Ile	ctt Leu	gga Gly	ggt Gly 820	Ala	gca Ala	tta Leu	act	Pro 825	aaa Lys	ttt Phe	gtg Val	aat Asn	gaa Glu 830	Asp	tgt Cys	2496	
agt Ser	cag Gln	ata Ile 835	tat	aaa Lys	ggt	aaa Lys	att Ile 840	Leu	tat Tyr	61A 888	aaa Lys	gat Asp 845	gct Ala	ttt	aca Thr	2544	
gat Asp	tta Leu 850	caa Gln	ttt Phe	atg Met	aat Asn	gac Asp 855	Tyr	atg Met	gat Asp	agt Ser	aaa Lys 860	aag Lys	aag Lys	ggc Gly	aat Asn	2592	
tgg Trp 865	tct Ser	aat Asn	gaa Glu	aat Asn	ggt Gly 870	ttt Phe	act Thr	aat Asn	act Thr	gat Asp 875	gat Asp	att Ile	caa Gln	att Ile	aaa Lys 880	2640	ş
Leu	Ala	Ser	Pro	Arg 885	Ser	Ser	Ala		Asp	Гув	Asn	Leu	Asn	Lys 895	Asn	2688	
Phe	Glu	Lув	Thr 900	Lys	Ser	Ile	Ģln	tta Leu 905	Ile	Glu	Asn	Phe	Asn 910	Arg	Ser	2736	
Asn	Phe	Val 915	Glu	Glu	Glu	Glu	Pro 920	ata Ile	Lys	Ala	Pro	Phe 925	Leu	Gly	Thr	2784	
Arg	Val 930	Leu	Gln	Ąsp	Ile	Glu 935	Ile	gac Asp	Phe	Asp	Lys 940	Leu	Ile	Phe	Tyr	2832	
Leu 945	Asp	Lys	Lys	Ala	Leu 950	Phe	Ser	ggt Gly	Gln	Trp 955	Gln	Ile	Lys	Lys	Asn 960	2880	
Lys	Gly	Gln	ser	Val 965	Glu	Glu	Tyr		Asn 970	Tyr	Leu	Asp	Ser	Tyr 975	Ala	2928	
aat Asn	Pro	Leu	Leu 980	Glu	Lys	Trp	Ile	Asn 985	Ile	Ile	Leu	Asp	Lys 990	Gly	Leu	2976	
att Ile	Ser	Pro 995	Lys	Val	Val	Tyr	Gly .000	Tyr	Phe	Arg	Сув	Gly 1005	Arg	Asn	Asp	3024	
	Ser 010	Ile	Tyr	Leu	Phe 1	Asp .015	Asn	Val	Ser	naA I	Lys 020	Arg	Ile	Ser	Glu	3072	
Phe 1	Asn	ttt Phe	cct Pro	aga Arg	caa Gln 1030	Lys	tcg Ser	gga Gly	aat Asn	aat Asn 1035	Leu	tgt Cys	att Ile	gca Ala	gat Asp 1040	3120	
ttt Phe '	tac Tyr	tgt Cys	Авр	ctt Leu 045	aaa Lys	aat Asn	aat Asn	qaA	cca Pro 050	gta Val	gat Asp	ata Ile	Phe	cca Pro 055	atg Met	3168	
caa (gca Ala	gta Val	aca Thr	atg Met	999 999	gaa Glu	ata Ile	gct Ala	agc Ser	gaa Glu	tat Tyr	tcc Ser	caa Gln	gaa Glu	tta Leu	3216	

32 1060 1065 1070 ttt aaa gct gat aaa tat agt gat tat tta ata ttt cat ggt tta acc 3264 Phe Lys Ala Asp Lys Tyr Ser Asp Tyr Leu Ile Phe His Gly Leu Thr 1075 1080 1085 gtt caa tta gca gaa gct ctt gca gaa tat gtt cat tca ata gta aga 3312 Val Gln Leu Ala Glu Ala Leu Ala Glu Tyr Val His Ser Ile Val Arg 1090 1095 1100 att gaa tgc gga ttt aaa tca tat gag cca aac aat aac cgt gat ata 3360 Ile Glu Cys Gly Phe Lys Ser Tyr Glu Pro Asn Asn Asn Arg Asp Ile 1105 1110 1115 tta gct caa aaa tat aga gga gct aga tac tca ttt ggt tat cca gct Leu Ala Gln Lys Tyr Arg Gly Ala Arg Tyr Ser Phe Gly Tyr Pro Ala 1130 tgt cct aaa gtt tct gat tca aat ata cag tta tca tta ttg gat aca 3456 Cys Pro Lys Val Ser Asp Ser Asn Ile Gln Leu Ser Leu Leu Asp Thr 1145 aaa agg att aat tta aca atg gat gaa tca gag caa tta cat cct gaa 3504 Lys Arg Ile Asn Leu Thr Met Asp Glu Ser Glu Gln Leu His Pro Glu 1160 1165 caa agt act act gct ata att tca ctt cat tca aaa gca aaa tat ttt 3552 Gln Ser Thr Thr Ala Ile Ile Ser Leu His Ser Lys Ala Lys Tyr Phe 1170 1175 1180 agt gcc taa 3561 Ser Ala 1185 <210> 8 <211> 1186 <212> PRT <213> Prochlorococcus marinus Met Val Ser Phe Arg Asn Tyr Leu Asn Arg Asp Asp Lys Pro Ile Ile 10 Ile Phe Asp Gly Gly Thr Gly Thr Ser Phe Gln Asn Leu Asn Leu Ser Ser His Asp Phe Gly Gly Asp Asp Leu Glu Gly Cys Asn Glu Asn Leu 40 45 Val Leu Ser Ser Pro Asn Thr Val Glu Gln Val His Asn Ser Phe Leu 55 Glu Ala Gly Cys His Val Ile Glu Thr Asn Thr Phe Gly Ala Ser Ser

Ile Val Leu Asp Glu Tyr Ser Ile Ser Asn Lys Ala Tyr Glu Ile Asn

Lys Lys Ala Ala Gln Ile Ala Lys Lys Cys Ala Asn Leu Phe Ser Ser

105

Ile Asn Thr Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr Lys
115 120 125

Leu Pro Thr Leu Gly His Ile Ser Phe Asp Lys Leu Lys Asp Ser Tyr 130 140

Glu Glu Gln Ile Asn Gly Leu Ile Asp Gly Gly Ile Asp Leu Leu Leu 145 150 155 160

Ile Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ser Ala Leu Ser Ala 165 170 175

Ser Gln Glu Val Ile Lys Asn Arg Asn Ile Glu Leu Pro Ile Met Ile 180 185 190

Ser Ile Thr Met Glu Thr Thr Gly Thr Met Leu Val Gly Ser Asp Ile 195 200 205

Ala Ser Ala Leu Thr Ile Leu Glu Pro Tyr Asn Ile Asp Ile Leu Gly
210 215 220

Leu Asn Cys Ala Thr Gly Pro Val Gln Met Lys Glu His Ile Lys Tyr 225 230 235 240

Leu Ala Glu Asn Ser Pro Phe Ala Ile Ser Cys Ile Pro Asn Ala Gly
245 250 255

Leu Pro Glu Asn Ile Gly Gly Val Ala His Tyr Lys Leu Thr Pro Leu 260 265 270

Glu Leu Lys Met Gln Leu Met Asn Phe Ile Tyr Asp Phe Asn Val Gln 275 280 285

Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Lys His Leu 290 295 300

Ser Ser Ile Ile Glu Glu Ile Val Asp Lys Lys Ile Asn Lys Arg Leu 305 310 315 320

Pro Thr Val Lys Thr Asn Phe Val Pro Ser Ala Ala Ser Ile Tyr Asn 325 330 335

Ala Val Pro Tyr Lys Gln Asp Asn Ser Ile Leu Ile Val Gly Glu Arg 340 345 350

Leu Asn Ala Ser Gly Ser Lys Lys Val Arg Glu Leu Leu Asn Glu Asp 355 360 365

Asp Trp Asp Gly Leu Leu Ser Ile Ala Lys Gln Gln Gln Lys Glu Asn 370 375 380

Ala His Ile Leu Asp Val Asp Val Asp Tyr Val Gly Arg Asp Gly Val
385 390 395 400

Lys Asp Met Lys Glu Ile Thr Ser Arg Leu Val Thr Asn Ile Asn Leu 405 410 415

Pro Leu Met Ile Asp Ser Thr Glu Ala Asp Lys Met Glu Ser Gly Leu
420 425 430

Lys Thr Val Gly Gly Lys Cys Ile Ile Asn Ser Thr Asn Tyr Glu Asp 435 440 445

- Gly Asp Asp Arg Phe Asn Gln Val Leu Arg Leu Ala Leu Asp Tyr Gly 450 460
- Ala Gly Ile Val Ile Gly Thr Ile Asp Glu Asp Gly Met Ala Arg Thr 465 470 475 480
- Ser Gln Lys Lys Tyr Asp Ile Ala Lys Arg Ala Leu Ile Lys Thr Arg
 485 490 490
- Ser Ser Gly Leu Ala Asp Tyr Glu Ile Phe Phe Asp Pro Leu Ala Leu 500 505 510
- Pro Ile Ser Thr Gly Ile Glu Glu Asp Arg Leu Asn Ala Lys Ala Thr 515 520 525
- Ile Glu Ala Ile Ser Lys Ile Arg Lys Ser Phe Pro Asp Ile His Ile 530 540
- Ile Leu Gly Ile Ser Asn Ile Ser Phe Gly Leu Ser Pro Leu Ser Arg
 555 555 560
- Ile Asn Leu Asn Ser Ile Phe Leu Asp Glu Cys Ile Lys Ala Gly Leu 565 570 575
- Asp Ser Ala Ile Ile Ala Pro Asn Lys Ile Leu Pro Leu Ser Lys Ile 580 585 590
- Ser Ala Glu Thr Lys Lys Leu Cys Leu Asp Leu Ile Tyr Asp Arg Arg 595 600 605
- Asn Phe Glu Asn Glu Ile Cys Ile Tyr Asp Pro Leu Val Glu Leu Thr 610 620
- Lys Ala Phe Gln Asp Ile Thr Ile Ser Asp Phe Lys Lys Gly Ser Thr 625 635 640
- Ser Asn Lys Asn Leu Thr Leu Glu Glu Lys Leu Lys Asn His Ile Val 645 650 655
- Asp Gly Glu Lys Ile Gly Leu Glu Glu Gln Leu Asn Asn Ala Leu Lys 660 665 670
- Lys Tyr Lys Pro Leu Glu Ile Ile Asn Thr Tyr Leu Leu Asp Gly Met 675 680 685
- Lys Val Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe 690 695 700
- Val Leu Gln Ser Ala Glu Thr Met Lys Phe Ala Val Ser Val Leu Glu 705 710 715 720
- Pro His Met Glu Thr Val Asp Glu Lys Ile Ser Asn Gly Lys Leu Leu
 725 730 735
- Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Leu Val
 740 745 750
- Asp Ile Ile Leu Ser Asn Asn Gly Phe Asp Val Ile Asn Leu Gly Ile 755 760 765
- Lys Gln Asp Val Ser Ala Ile Ile Asp Ala Gln Lys Lys His Lys Ala

WO 03/087386 35 775 780 Asp Cys Ile Ala Met Ser Gly Leu Leu Val Lys Ser Thr Ala Phe Met 790 795 Lys Asp Asn Leu Glu Ala Phe Asn Asn Ala Glu Ile Asn Val Pro Val 805 810 Ile Leu Gly Gly Ala Ala Leu Thr Pro Lys Phe Val Asn Glu Asp Cys 825 Ser Gln Ile Tyr Lys Gly Lys Ile Leu Tyr Gly Lys Asp Ala Phe Thr Asp Leu Gln Phe Met Asn Asp Tyr Met Asp Ser Lys Lys Gly Asn 850 855 Trp Ser Asn Glu Asn Gly Phe Thr Asn Thr Asp Asp Ile Gln Ile Lys 870 Leu Ala Ser Pro Arg Ser Ser Ala Lys Asp Lys Asn Leu Asn Lys Asn 885 890 Phe Glu Lys Thr Lys Ser Ile Gln Leu Ile Glu Asn Phe Asn Arg Ser 905 Asn Phe Val Glu Glu Glu Pro Ile Lys Ala Pro Phe Leu Gly Thr 920 925

Arg Val Leu Gln Asp Ile Glu Ile Asp Phe Asp Lys Leu Ile Phe Tyr

Leu Asp Lys Lys Ala Leu Phe Ser Gly Gln Trp Gln Ile Lys Lys Asn

Lys Gly Gln Ser Val Glu Glu Tyr Asn Asn Tyr Leu Asp Ser Tyr Ala 970

Asn Pro Leu Leu Glu Lys Trp Ile Asn Ile Ile Leu Asp Lys Gly Leu 985

Ile Ser Pro Lys Val Val Tyr Gly Tyr Phe Arg Cys Gly Arg Asn Asp 1000

Asn Ser Ile Tyr Leu Phe Asp Asn Val Ser Asn Lys Arg Ile Ser Glu 1015

Phe Asn Phe Pro Arg Gln Lys Ser Gly Asn Asn Leu Cys Ile Ala Asp 1035

Phe Tyr Cys Asp Leu Lys Asn Asn Asp Pro Val Asp Ile Phe Pro Met 1050

Gln Ala Val Thr Met Gly Glu Ile Ala Ser Glu Tyr Ser Gln Glu Leu 1065

Phe Lys Ala Asp Lys Tyr Ser Asp Tyr Leu Ile Phe His Gly Leu Thr 1075 1080

Val Gln Leu Ala Glu Ala Leu Ala Glu Tyr Val His Ser Ile Val Arg 1090 1095 1100

WO 03/087386 PCT/EP03/04010

Ile Glu Cys Gly Phe Lys Ser Tyr Glu Pro Asn Asn Asn Arg Asp Ile
1105 1110 1115 1120

Leu Ala Gln Lys Tyr Arg Gly Ala Arg Tyr Ser Phe Gly Tyr Pro Ala
1125 1130 1135

Cys Pro Lys Val Ser Asp Ser Asn Ile Gln Leu Ser Leu Leu Asp Thr
1140 1145 1150

Lys Arg Ile Asn Leu Thr Met Asp Glu Ser Glu Gln Leu His Pro Glu 1155 1160 1165

Gln Ser Thr Thr Ala Ile Ile Ser Leu His Ser Lys Ala Lys Tyr Phe 1170 1175 1180

Ser Ala 1185

<210> 9 <211> 3048

<212> DNA

<213> Thermus thermophilus

<220>

<221> CDS

<222> (1)..(3045)

<223> RTT00266

<400> 9

atg cgg gcc tac aag gag gcg gca cgg ggg ctt ctt aag ggc ggg gtg 48 Met Arg Ala Tyr Lys Glu Ala Ala Arg Gly Leu Leu Lys Gly Gly Val 1 5 10 15

gac ctc atc ctc ttg gag acc gcc cag gac atc ctc cag gtg cgc tgc 96
Asp Leu Ile Leu Leu Glu Thr Ala Gln Asp Ile Leu Gln Val Arg Cys
20 25 30

gcc gtc ttg gcg gtg cgg gag gcc atg gcc gag gtg ggc cgg gag gtg 144 Ala Val Leu Ala Val Arg Glu Ala Met Ala Glu Val Gly Arg Glu Val 35 40

ccc ctc cag gtc cag gtg acc ttt gag gcc acg ggg acg atg ctc gtg
Pro Leu Gln Val Gln Val Thr Phe Glu Ala Thr Gly Thr Met Leu Val
50 55 60

ggc acg gac gag cag gcc ctg gcc gct ctg gag agc ctc ccc gtg
Gly Thr Asp Glu Gln Ala Ala Leu Ala Ala Leu Glu Ser Leu Pro Val
65 70 75 80

gac gtg gtg ggg atg aac tgc gcc acg ggc ccc gac ctc atg gac agc 288 Asp Val Val Gly Met Asn Cys Ala Thr Gly Pro Asp Leu Met Asp Ser 85 90

aag gtg cgc tac ttc gcc gag cac agc acc cgc ttc gtc tcc tgc ctc 336 Lys Val Arg Tyr Phe Ala Glu His Ser Thr Arg Phe Val Ser Cys Leu 100 105 110

ccg aac gcg ggc ctg ccc cgg aac gag ggg ggg agg gtg gtc tac gac 384 Pro Asn Ala Gly Leu Pro Arg Asn Glu Gly Gly Arg Val Val Tyr Asp 115 120 125

												٠,				
									3	37						
					ctc Leu	_	_				_			_		432
					gtg Val 150											480
					gag Glu											528
					ccc Pro											576
					agc Ser											624
					ttc Phe											672
ggc Gly 225	atc Ile	ctc Leu	gcc Ala	ctc Leu	gcc Ala 230	cgg Arg	gag Glu	cag Gln	gtg Val	gag Glu 235	gag Glu	99 9 91y	gcc Ala	cac His	gcc Ala 240	720
_	_				gcc Ala		_			_				_		768
					cat His											816
					cct Pro											864
		Arg			ctg Leu		Ser	Ala	Asn		Glu	Asp				912
cgc Arg 305	ttt Phe	gac Asp	cgg Arg	gtg Val	gcc Ala 310	tcc Ser	ctg Leu	gcc Ala	aag Lys	gcc Ala 315	cac His	Gly 999	gcg Ala	gcc Ala	ctc Leu 320	960
					gac Asp											1008
aag Lys	gtg Val	cgg Arg	gtg Val 340	gcc Ala	ctg Leu	agg Arg	atg Met	tac Tyr 345	gag Glu	cgc Arg	ctc Leu	acg Thr	gag Glu 350	cac His	cac His	1056
					gac Asp											1104
					gag Glu											1152

	370)				375				38	380					
acc			. aac	cto	cgg			ctt		~~~		~~~				1000
Ala 385	Ile	Glu	Glu	Leu	Arg 390	Glu	Arg	Leu	Pro	999 Gly 395	Val	Gly	Phe	Val	Leu 400	1200
cgg Arg	gtc Val	tcc Ser	aac Asn	gto Val 405	tcc Ser	ttc Phe	999 Gly	ctc Leu	aag Lys 410	Pro	cgg Arg	gcg Ala	agg Arg	cgc Arg 415	gtc Val	1248
ctg Leu	aac Asn	tcc Ser	gtc Val 420	Phe	ctg Leu	gac Asp	gag Glu	gcg Ala 425	agg Arg	aaa Lys	cgg Arg	ggc Gly	ctc Leu 430	acc Thr	gcg Ala	1296
gcc Ala	atc Ile	gtg Val 435	gac	gcg Ala	999 999	aag Lys	atc Ile 440	ctc Leu	ccc Pro	ata Ile	agc Ser	cag Gln 445	atc Ile	CCC Pro	ga g Glu	1344
gag Glu	gcc Ala 450	Tyr	gcc Ala	ctc Leu	gcc Ala	tta Leu 455	gac	ctc Leu	atc Ile	tac Tyr	gac Asp 460	cgc Arg	cgc Arg	aag Lys	gag Glu	1392
ggc Gly 465	ttt Phe	gac Asp	ccc Pro	ctc Leu	ctc Leu 470	gcc Ala	ttc Phe	atg Met	gcc Ala	tac Tyr 475	ttt Phe	gag Glu	gcc Ala	cac His	aag Lys 480	1440
gag Glu	gac Asp	ccg Pro	99 9 91y	aag Lys 485	agg Arg	gag Glu	gac Asp	gcc Ala	ttc Phe 490	ctg Leu	gcc Ala	ctt Leu	ccc Pro	ctt Leu 495	ctg Leu	1488
gag Glu	agg Arg	ctc Leu	aag Lys 500	cgc Arg	cgc Arg	gtg Val	gtg Val	gag Glu 505	Gly 99 9	agg Arg	aag Lys	cag Gln	ggc Gly 510	ctc Leu	gag Glu	1536
gcc Ala	gac Asp	ctg Leu 515	gag Glu	gag Glu	gcc Ala	ctg Leu	aag Lys 520	gcg Ala	999 Gly	cac His	aag Lys	ccc Pro 525	ttg Leu	gac Asp	ctc Leu	1584
Ile	aac Asn 530	gly	ccc Pro	ctc Leu	ctc Leu	gcg Ala 535	ggg Gly	atg Met	aag Lys	gag Glu	gtg Val 540	61À 333	gac Asp	ctc Leu	ttc Phe	1632
999 Gly 545	gcg Ala	999 Gly	aag Lys	atg Met	cag Gln 550	ctc Leu	ccc Pro	ttc Phe	gtc Val	ctc Leu 555	cag Gln	gcc Ala	gcc Ala	gag Glu	gtg Val 560	1680
atg Met	aag Lys	cgg Arg	gcg Ala	gtg Val 565	gcc Ala	tac Tyr	ctc Leu	gag Glu	ccc Pro 570	cac His	atg Met	gag Glu	aag Lys	aag Lys 575	G1y 999	1728
gag Glu	ggc Gly	Lys	ggt Gly 580	acc Thr	ctg Leu	gtc Val	ctc Leu	gcc Ala 585	acc Thr	gtc Val	aag Lys	G1Y 999	gac Asp 590	gtg Val	cac His	1776
gac Asp	Ile	ggc Gly 595	aag Lys	aac Asn	ctg Leu	gtg Val	gac Asp 600	atc Ile	atc Ile	ctc Leu	agc Ser	aac Asn 605	aac Asn	ggc Gly	tac Tyr	1824
cgg (gtg Val 610	gtg Val	aac Asn	ctg Leu	Gly	atc Ile 615	aag Lys	gtg Val	ccc Pro	Ile	gag Glu 620	gag Glu	atc Ile	ctg Leu	aag Lys	1872

													•				
											39						
A	1a 25	gtg Val	gag Glu	gcg Ala	Cac His	aag Lys 630	Pro	cac His	gcc Ala	gtg Val	ggc Gly 635	atg Met	tcg Ser	ggc Gly	Leu	ctg Leu 640	1920
9 V	tg al	aag Lys	agc Ser	acc Thr	ctg Leu 645	gtg Val	atg Met	aag Lys	gag Glu	aac Asn 650	ctg Leu	gag Glu	tac Tyr	atg Met	cgg Arg 655	gat Asp	1968
a A	gg .rg	ggc Gly	tac Tyr	acc Thr 660	ctc Leu	ccc Pro	gtg Val	atc Ile	ctg Leu 665	ggc	ggg Gly	gcc Ala	gcc Ala	ctc Leu 670	acc Thr	cgg Arg	2016
						ctt Leu											2064
						ggc Gly											2112
7 7	1a 05	Pro	Pro	Glu	Leu	acc Thr 710	Arg	Lys	Ala	Pro	Ala 715	Arg	Pro	Lys	Arg	Glu 720	2160
A	la.	Pro	ГÀв	Val	Ala 725	Pro	Arg	Ala	Arg	Pro 730	Val	Gly	Glu	Ala	Pro 735	Ala	2208
V	al	Pro	Arg	Pro 740	Pro	ttc Phe	Phe	Gly	Val 745	Arg	Val	Glu	Glu	Gly 750	Leu	Asp	2256
						cac His											2304
G	ln	Trp 770	Gly	Tyr	Ser	cgc Arg	Lys 775	Gly	Phe	Pro	Gly	Arg 780	Arg	Gly	Arg	Pro	2352
T 7	rp 85	Trp	Ser	Gly	Arg	cgg Arg 790	Ser	Leu	Ser	Ser	Arg 795	Gly	Ser	Ser	Arg	Arg 800	2400
A	rg	Trp	Arg	Lys	Gly 805	ggc	Leu	Asn	Pro	Arg 810	Ser	Ser	Thr	Ala	Ser 815	Ser	2448
P	ro	Trp	Pro	Gly 820	Arg	gga Gly	Gly	Ala	Ser 825	Arg	Leu	Leu	Pro	Arg 830	Asp	Gly	2496
G	ly	Gly	Ala 835	Gly	Ala	ctt Leu	Pro	Leu 840	Pro	Pro	Āla	Lys	Gly 845	Arg	Gly	Pro	2544
G	lu	Pro 850	Arg	Gly	Leu	ctt Leu	Pro 855	Pro	Pro	Val	Сув	Arg 860	Ala	Phe	Gly	Gly	2592
C A	ga rg	ggc Gly	gga Gly	ctg Leu	gat Asp	gcc Ala	caa Gln	gga Gly	ggc Gly	ctt Leu	ccg Pro	ggc Gly	999 Gly	ggc Gly	cgg Arg	gac Asp	2640

40 870 865 875 880 gtc ctc ggg gtc cag ctc gtc acc atg ggg gag gcc cct tcc cga aag 2688 Val Leu Gly Val Gln Leu Val Thr Met Gly Glu Ala Pro Ser Arg Lys 890 gcc cag gcc ctc ttt gcg tcc ggg gcc tac cag gac tac ctc ttc gtc 2736 Ala Gln Ala Leu Phe Ala Ser Gly Ala Tyr Gln Asp Tyr Leu Phe Val 900 905 cac ggc ttc agc gtg gag atg acc gag gcc ttg gcg gag tac tgg cac 2784 His Gly Phe Ser Val Glu Met Thr Glu Ala Leu Ala Glu Tyr Trp His 920 aag agg atg egg eag atg tgg gge ate gee eac aag gae gee ace gag 2832 Lys Arg Met Arg Gln Met Trp Gly Ile Ala His Lys Asp Ala Thr Glu 935 940 atc cag aag ctc ttc cag cag ggc tac cag ggg gcc cgc tac tcc ttc 2880 Ile Gln Lys Leu Phe Gln Gln Gly Tyr Gln Gly Ala Arg Tyr Ser Phe 950 955 ggc tac ccc gcc tgc ccg gac ctc gcc gac cag gcc aag ctg gac cgg 2928 Gly Tyr Pro Ala Cys Pro Asp Leu Ala Asp Gln Ala Lys Leu Asp Arg 965 970 ctc atg ggc ttc cac cgg gtg ggg gtg cac ctc acg gag aac ttc cag 2976 Leu Met Gly Phe His Arg Val Gly Val His Leu Thr Glu Asn Phe Gln 985 ctg gag ccg gag cac gcc acc agc gcc ctc gtg gtc cac cac ccc gag 3024 Leu Glu Pro Glu His Ala Thr Ser Ala Leu Val Val His His Pro Glu 1000 995 gcc cgc tac ttc agc gtg gac tag 3048 Ala Arg Tyr Phe Ser Val Asp 1010 <210> 10 <211> 1015 <212> PRT

<213> Thermus thermophilus

Met Arg Ala Tyr Lys Glu Ala Ala Arg Gly Leu Leu Lys Gly Gly Val

Asp Leu Ile Leu Leu Glu Thr Ala Gln Asp Ile Leu Gln Val Arg Cys 25

Ala Val Leu Ala Val Arg Glu Ala Met Ala Glu Val Gly Arg Glu Val 35

Pro Leu Gln Val Gln Val Thr Phe Glu Ala Thr Gly Thr Met Leu Val 55 60

Gly Thr Asp Glu Gln Ala Ala Leu Ala Ala Leu Glu Ser Leu Pro Val

Asp Val Val Gly Met Asn Cys Ala Thr Gly Pro Asp Leu Met Asp Ser 90

- Lys Val Arg Tyr Phe Ala Glu His Ser Thr Arg Phe Val Ser Cys Leu 100 105 110
- Pro Asn Ala Gly Leu Pro Arg Asn Glu Gly Gly Arg Val Val Tyr Asp 115 120 125
- Leu Thr Pro Glu Glu Leu Ala Lys Trp His Leu Lys Phe Val Ala Glu 130 135 140
- Tyr Gly Val Asn Ala Val Gly Gly Cys Cys Gly Thr Gly Pro Glu His 145 150 155 160
- Ile Arg Lys Val Ala Glu Ala Val Lys Gly Leu Ala Pro Lys Pro Arg 165 170 175
- Pro Glu Ser Phe Pro Pro Gln Val Ala Ser Leu Tyr Gln Ala Val Ser 180 185 190
- Leu Lys Gln Glu Ala Ser Leu Phe Leu Val Gly Glu Arg Leu Asn Ala 195 200 205
- Thr Gly Ser Lys Arg Phe Arg Glu Met Leu Phe Ala Arg Asp Leu Glu 210 215 220
- Gly Ile Leu Ala Leu Ala Arg Glu Gln Val Glu Glu Gly Ala His Ala 225 230 235 240
- Leu Asp Leu Ser Val Ala Trp Thr Gly Arg Asp Glu Leu Glu Asp Leu 245 250 255
- Arg Trp Leu Leu Pro His Leu Ala Thr Ala Leu Thr Val Pro Val Met 260 265 270
- Val Asp Ser Thr Ser Pro Glu Ala Met Glu Leu Ala Leu Lys Tyr Leu 275 280 285
- Pro Gly Arg Val Leu Leu Asn Ser Ala Asn Leu Glu Asp Gly Leu Glu 290 295 300
- Arg Phe Asp Arg Val Ala Ser Leu Ala Lys Ala His Gly Ala Ala Leu 305 310 315 320
- Val Val Leu Ala Ile Asp Glu Lys Gly Met Ala Lys Thr Arg Glu Glu 325 330 335
- Lys Val Arg Val Ala Leu Arg Met Tyr Glu Arg Leu Thr Glu His His 340 345 350
- Gly Leu Arg Pro Glu Asp Leu Leu Phe Asp Leu Leu Thr Phe Pro Ile 355 360 365
- Thr Gln Gly Asp Glu Glu Ser Arg Pro Leu Ala Lys Glu Thr Leu Leu 370 375 380
- Ala Ile Glu Glu Leu Arg Glu Arg Leu Pro Gly Val Gly Phe Val Leu 385 390 395 400
- Arg Val Ser Asn Val Ser Phe Gly Leu Lys Pro Arg Ala Arg Arg Val 405 410 415
- Leu Asn Ser Val Phe Leu Asp Glu Ala Arg Lys Arg Gly Leu Thr Ala

420 425 430

Ala Ile Val Asp Ala Gly Lys Ile Leu Pro Ile Ser Gln Ile Pro Glu
435 440 445

Glu Ala Tyr Ala Leu Ala Leu Asp Leu Ile Tyr Asp Arg Arg Lys Glu
450 460

Gly Phe Asp Pro Leu Leu Ala Phe Met Ala Tyr Phe Glu Ala His Lys 465 470 475 480

Glu Asp Pro Gly Lys Arg Glu Asp Ala Phe Leu Ala Leu Pro Leu Leu 485 490 495

Glu Arg Leu Lys Arg Arg Val Val Glu Gly Arg Lys Gln Gly Leu Glu 500 505 510

Ala Asp Leu Glu Glu Ala Leu Lys Ala Gly His Lys Pro Leu Asp Leu
515 520 525

Ile Asn Gly Pro Leu Leu Ala Gly Met Lys Glu Val Gly Asp Leu Phe 530 540

Gly Ala Gly Lys Met Gln Leu Pro Phe Val Leu Gln Ala Ala Glu Val 545 550 555 560

Met Lys Arg Ala Val Ala Tyr Leu Glu Pro His Met Glu Lys Lys Gly
565 570 575

Glu Gly Lys Gly Thr Leu Val Leu Ala Thr Val Lys Gly Asp Val His 580 585 590

Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly Tyr
595 600 605

Arg Val Val Asn Leu Gly Ile Lys Val Pro Ile Glu Glu Ile Leu Lys 610 615 620

Ala Val Glu Ala His Lys Pro His Ala Val Gly Met Ser Gly Leu Leu 625 630 635 640

Val Lys Ser Thr Leu Val Met Lys Glu Asn Leu Glu Tyr Met Arg Asp 645 650 655

Arg Gly Tyr Thr Leu Pro Val Ile Leu Gly Gly Ala Ala Leu Thr Arg
660 665 670

Ser Tyr Val Glu Glu Leu Lys Ala Ile Tyr Pro Asn Val Tyr Tyr Ala 675 680 685

Glu Asp Ala Phe Glu Gly Leu Arg Leu Met Glu Glu Leu Thr Gly His 690 695 700

Ala Pro Pro Glu Leu Thr Arg Lys Ala Pro Ala Arg Pro Lys Arg Glu 705 710 715 720

Ala Pro Lys Val Ala Pro Arg Ala Arg Pro Val Gly Glu Ala Pro Ala
725 730 735

Val Pro Arg Pro Pro Phe Phe Gly Val Arg Val Glu Glu Gly Leu Asp
740 745 750

Leu Ala Thr Ile Ala His Tyr Val Asn Lys Leu Ala Leu Tyr Arg Gly 755 760 765

Gln Trp Gly Tyr Ser Arg Lys Gly Phe Pro Gly Arg Arg Gly Arg Pro
770 780

Trp Trp Ser Gly Arg Arg Ser Leu Ser Ser Arg Gly Ser Ser Arg Arg 785 790 795 800

Arg Trp Arg Lys Gly Gly Leu Asn Pro Arg Ser Ser Thr Ala Ser Ser 805 810 815

Pro Trp Pro Gly Arg Gly Gly Ala Ser Arg Leu Leu Pro Arg Asp Gly 820 825 830

Gly Gly Ala Gly Ala Leu Pro Leu Pro Pro Ala Lys Gly Arg Gly Pro 835 840 845

Glu Pro Arg Gly Leu Leu Pro Pro Pro Val Cys Arg Ala Phe Gly Gly 850 855 860

Arg Gly Gly Leu Asp Ala Gln Gly Gly Leu Pro Gly Gly Gly Arg Asp 880

Val Leu Gly Val Gln Leu Val Thr Met Gly Glu Ala Pro Ser Arg Lys 885 890 895

Ala Gln Ala Leu Phe Ala Ser Gly Ala Tyr Gln Asp Tyr Leu Phe Val

His Gly Phe Ser Val Glu Met Thr Glu Ala Leu Ala Glu Tyr Trp His 915 920 925

Lys Arg Met Arg Gln Met Trp Gly Ile Ala His Lys Asp Ala Thr Glu 930 935 940

Ile Gln Lys Leu Phe Gln Gln Gly Tyr Gln Gly Ala Arg Tyr Ser Phe 945 950 955 960

Gly Tyr Pro Ala Cys Pro Asp Leu Ala Asp Gln Ala Lys Leu Asp Arg 965 970 975

Leu Met Gly Phe His Arg Val Gly Val His Leu Thr Glu Asn Phe Gln 980 985 990

Leu Glu Pro Glu His Ala Thr Ser Ala Leu Val Val His His Pro Glu 995 1000 1005

Ala Arg Tyr Phe Ser Val Asp 1010 1015

<210> 11

<211> 3441

<212> DNA

<213> Bacillus halodurans

<220>

<221> CDS

<222> (1)..(3438)

<223> RHD05550

-20	· /	-														
atg Met 1	Thr	aaa Lys	tcg Ser	ttg Leu 5	Phe	gaa Glu	caa Gln	Cag Gln	tta Leu 10	gag Glu	cga Arg	aaa Lys	atc Ile	gtc Val 15	atc Ile	48
ctt Leu	gat Asp	999 999	gcg Ala 20	Met	Gly	acc Thr	atg Met	tta Leu 25	Gln	gcc Ala	gcg Ala	aat Asn	cta Leu 30	acc Thr	gct Ala	96
gat Asp	gac Asp	ttt Phe 35	ggc	gga Gly	gaa Glu	gag Glu	tat Tyr 40	gaa Glu	G1 y 333	tgt Cys	aat Asn	gaa Glu 45	tat Tyr	tta Leu	aat Asn	144
gag Glu	acg Thr 50	Ala	ccc Pro	cat His	gtc Val	gtt Val 55	Glu	gac Asp	att Ile	cat His	cgc Arg 60	gca Ala	tac Tyr	tta Leu	gag Glu	192
gca Ala 65	gga Gly	gca Ala	gac Asp	gtc Val	att Ile 70	Ala	acg Thr	aac Asn	acg Thr	ttc Phe 75	999 Gly	gca Ala	aca Thr	gat Asp	atc Ile 80	240
gtt Val	ctt Leu	gac Asp	gat Asp	tat Tyr 85	gat Asp	ctc Leu	gga Gly	tac Tyr	aaa Lys 90	gca Ala	gag Glu	gag Glu	tta Leu	aac Asn 95	ata Ile	288
tgc Cys	gcg Ala	gtg Val	aaa Lys 100	atc Ile	gct Ala	aaa Lys	cgt Arg	gta Val 105	gct Ala	gaa Glu	gag Glu	ttt Phe	tcc Ser 110	act Thr	cca Pro	336
gat Asp	tgg Trp	cct Pro 115	cga Arg	ttc Phe	gtt Val	gca Ala	999 Gly 120	gcg Ala	atg Met	61Å aaa	ccg Pro	acg Thr 125	acg Thr	aaa Lys	tct Ser	384
Leu	Ser 130	Val	aca Thr	Gly	Gly	Ala 135	Thr	Phe	Glu	Gln	Leu 140	Ile	Glu	Ser	Tyr	432
cgc Arg 145	cag Gln	caa Gln	gct Ala	aca Thr	ggt Gly 150	cta Leu	att Ile	aaa Lys	ggc Gly	999 Gly 155	gcg Ala	gat Asp	att Ile	tta Leu	tta Leu 160	480
ctc Leu	gaa Glu	acg Thr	agc Ser	cag Gln 165	gat Asp	atg Met	cga Arg	aac Asn	gtg Val 170	aag Lys	gcg Ala	gct Ala	tat Tyr	tta Leu 175	gga Gly	528
ctg Leu	agc Ser	caa Gln	gcg Ala 180	caa Gln	aaa Lys	gag Glu	cta Leu	gag Glu 185	gtg Val	aaa Lys	ctg Leu	cct Pro	ctc Leu 190	att Ile	att Ile	576
er Ser	gga Gly	acg Thr 195	att Ile	gaa Glu	ccg Pro	atg Met	gga Gly 200	aca Thr	acg Thr	ctc Leu	gcc Ala	ggc Gly 205	caa Gln	aac Asn	atc Ile	624
3lu	gcg Ala 210	ttc Phe	tat Tyr	ttg Leu	tca Ser	tta Leu 215	gag Glu	cat His	atg Met	aat Asn	ccc Pro 220	gtc Val	gtt Val	gtc Val	ggt Gly	672
etc Leu 225	aac Asn	tgc Cys	gct Ala	Thr	gga Gly 230	cca Pro	gaa Glu	ttt Phe	Met	cgc Arg 235	gat Asp	cac His	ctc Leu	cgt Arg	tct Ser 240	720
tt	tca	gac	ctt	gcg	acc	tgc	tct	gta	agc	tqt	tat	cca	aat	act	aaa	768

										45						
Leu	Ser	Asp	Leu	Ala 245		Сув	Ser	Val			Tyr	Pro		Ala 255	Gly	
								cac His 265								816
gcc Ala	aag Lys	ctc Leu 275	gca Ala	ggt Gly	ttt Phe	gcg Ala	gaa Glu 280	aag Lys	ggc	tgg Trp	ttg Leu	aat Asn 285	atg Met	gtt Val	ggt Gly	864
ggc Gly	tgt Cys 290	tgc Cys	61 y 999	acg Thr	act Thr	cca Pro 295	gac Asp	cac His	att Ile	cgt Arg	gct Ala 300	ctt Leu	ttg Leu	gac Asp	gtt Val	912
atg Met 305	aag Lys	caa Gln	ttt Phe	gag Glu	ccg Pro 310	Arg	caa Gln	cca Pro	aaa Lys	999 Gly 315	gat Asp	cac His	ccc Pro	cac His	tcg Ser 320	960
gtc Val	tca Ser	gga Gly	att Ile	gag Glu 325	cca Pro	ctg Leu	tta Leu	tac Tyr	gat Asp 330	gac Asp	agc Ser	atg Met	cgt Arg	cca Pro 335	cta Leu	1008
Phe	gtc Val	ggt Gly	gaa Glu 340	cgg Arg	aca Thr	aac Asn	gtc Val	atc Ile 345	61Å 888	tct Ser	cgt Arg	aaa Lys	ttt Phe 350	Lys	cgg Arg	1056
								gaa Glu								1104
caa Gln	gtg Val 370	aag Lys	aaa Lys	Gly 999	gcc Ala	cac His 375	gtt Val	atc Ile	gat Asp	gtt Val	tgt Cys 380	ctt Leu	gct Ala	gat Asp	ccg Pro	1152
Asp 385	Arg	Aep	Glu	Met	Glu 390	Asp	Met	Glu	Glu	Phe 395	Leu	Lys	Phe	Val	400	1200
aac Asn	aaa Lys	gtg Val	aag Lys	gta Val 405	ccg Pro	ctc Leu	atg Met	att Ile	gac Asp 410	tcc Ser	acc Thr	gac Asp	gaa Glu	aag Lys 415	gta Val	1248
att Ile	gaa Glu	caa Gln	gcg Ala 420	ctt Leu	acg Thr	tat Tyr	tca Ser	caa Gln 425	61 y 999	aaa Lys	gcg Ala	atc Ile	att Ile 430	aat Asn	tcg Ser	1296
atc Ile	aac Asn	tta Leu 435	gag Glu	gac Asp	ggc Gly	gaa Glu	gaa Glu 440	cgt Arg	ttt Phe	gaa Glu	aaa Lys	gtg Val 445	gtc Val	ccg Pro	ctc Leu	1344
gtc Val	cat His 450	aag Lys	tat Tyr	gga Gly	gcc Ala	gcg Ala 455	gtt Val	gtc Val	gtt Val	ggt Gly	acg Thr 460	atc Ile	gac Asp	gaa Glu	gaa Glu	1392
gga Gly 465	atg Met	gcg Ala	att Ile	Thr	gca Ala 470	gaa Glu	aaa Lys	aaa Lys	Leu	gcg Ala 475	gtt Val	gcg Ala	aaa Lys	cga Arg	tca Ser 480	1440
tac Tyr	Asp gac	ctg Leu	Leu	gta Val 485	aac Asn	aaa Lys	tac Tyr	aac Asn	att Ile 490	cgt Arg	ccg Pro	agc Ser	gat Asp	att Ile 495	att Ile	1488

						cca Pro										1536
						gtg Val										1584
						att Ile 535										1632
						gag Glu										1680
						gat Asp										1728
						tct Ser										1776
ctc Leu	tta Leu	ttt Phe 595	gaa Glu	acg Thr	aca Thr	gat Asp	gaa Glu 600	acg Thr	ctc Leu	gct Ala	gag Glu	ttc Phe 605	acc Thr	gcc Ala	ttt Phe	1824
						gag Glu 615										1872
ttg Leu 625	gaa Glu	gag Glu	cgg Arg	ttg Leu	gca Ala 630	aac Asn	tac Tyr	att Ile	gtt Val	gaa Glu 635	G1y 999	tca Ser	aag Lys	gac Asp	gga Gly 640	1920
ctg Leu	aca Thr	gaa Glu	gat Asp	tta Leu 645	gat Asp	aaa Lys	gcg Ala	ctc Leu	gcg Ala 650	aaa Lys	tat Tyr	gat Asp	gat Asp	ccg Pro 655	ctt Leu	1968
-					_	ctc Leu	_			_	-	_	•		-	2016
ttg Leu	ttt Phe	aac Asn 675	aat Asn	aac Asn	gag Glu	ctt Leu	att Ile 680	gtc Val	gct Ala	gaa Glu	gta Val	ttg Leu 685	caa Gln	agc Ser	gct Ala	2064
gag Glu	gtt Val 690	atg Met	aag Lys	gct Ala	tcc Ser	gtc Val 695	gcc Ala	cac His	ctt Leu	gag Glu	cca Pro 700	cat His	atg Met	gaa Glu	aag Lys	2112
aaa Lys 705	gca Ala	gac Asp	gat Asp	cat His	gga Gly 710	aaa Lys	gga Gly	aaa Lys	atc Ile	att Ile 715	ctt Leu	gcc Ala	acg Thr	gtc Val	aag Lys 720	2160
ggc Gly	gat Asp	gtt Val	cac His	gat Asp 725	atc Ile	ggg Gly	aaa Lys	aat Asn	cta Leu 730	gtg Val	gaa Glu	att Ile	att Ile	ttg Leu 735	agc Ser	2208
aat	aat	ggt	ttc	cgc	atc	gtg	aac	cta	gga	att	aaa	gtt	acc	tct	aat	2256

										47						
Asn	Asn	Gly	740	Arg	Ile	· Val	Asn	Leu 745		Ile	Lys	Val	750		Asn	
gag Glu	ctg Leu	att Ile 755	Glu	gcg	gtg Val	gcg Ala	aga Arg 760	Glu	aat Asn	cca Pro	gat Asp	gcg Ala 765	Ile	ggc	ttg Leu	2304
tca Ser	999 Gly 770	Leu	ctc Leu	gtc Val	aaa Lys	tca Ser 775	Ala	caa Gln	caa Gln	atg Met	gta Val 780	Leu	acc Thr	gcc Ala	caa Gln	2352
gat Asp 785	Leu	aag Lys	caa Gln	caa Gln	caa Gln 790	Ile	tcc Ser	att Ile	ccg Pro	att Ile 795	tta Leu	gtc Val	gga Gly	ggc	gca Ala 800	2400
gcc Ala	ctt Leu	acg Thr	cgg	aaa Lys 805	ttt Phe	acg Thr	aat Asn	aca Thr	aaa Lys 810	atc Ile	gct Ala	cca Pro	gag Glu	tat Tyr 815	gat Asp	2448
ggt Gly	ctc Leu	gtc Val	gtc Val 820	tac Tyr	gcg Ala	aag Lys	gat Asp	gcg Ala 825	atg Met	aac Asn	Gly 999	tta Leu	gag Glu 830	ctt Leu	gcc Ala	2496
aat Asn	aaa Lys	tta Leu 835	atg Met	aaa Lys	cct Pro	gat Asp	gaa Glu 840	cga Arg	gaa Glu	aag Lys	cta Leu	gcg Ala 845	gtc Val	tcc Ser	ctc Leu	2544
cat His	gaa Glu 850	gcg Ala	aag Lys	gag Glu	cag Gln	gcg Ala 855	aac Asn	tcg Ser	agg Arg	aca Thr	caa Gln 860	atg Met	gga Gly	gga Gly	Gly	2592
gga Gly 865	act Thr	gca Ala	gtt Val	gcg Ala	gta Val 870	aag Lys	ccg Pro	act Thr	cga Arg	tcc Ser 875	cat His	gtt Val	tcg Ser	aca Thr	acg Thr 880	2640
gtg Val	cct Pro	gta Val	gcg Ala	gtc Val 885	cca Pro	cct Pro	gat Asp	gtg Val	aag Lys 890	ccg Pro	cac His	att Ile	ttg Leu	cgc Arg 895	cac His	2688
cat His	agc Ser	att Ile	gcc Ala 900	cat His	tta Leu	gag Glu	ccg Pro	tat Tyr 905	att Ile	aac Asn	atg Met	cag Gln	atg Met 910	ttg Leu	tta Leu	2736
gga Gly	cgt Arg	cac His 915	tta Leu	ggc Gly	tta Leu	caa Gln	999 Gly 920	ааа Lув	gtg Val	agc Ser	cgc Arg	ctg Leu 925	ctt Leu	gca Ala	gaa Glu	2784
aaa Lys	gac Asp 930	gag Glu	aag Lys	gct Ala	ctt Leu	gaa Glu 935	tta Leu	aaa Lys	gaa Glu	aaa Lys	gtt Val 940	gat Asp	gcg Ala	cta Leu	ctc Leu	2832
acc Thr 945	agg Arg	gtg Val	aaa Lys	gag Glu	gag Glu 950	cag Gln	ctc Leu	atg Met	Glu	gcc Ala 955	cat His	ggc Gly	atg Met	tat Tyr	cag Gln 960	2880
ttt Phe	ttt Phe	cct Pro	gcc Ala	cag Gln 965	tcg Ser	gat Asp	G1 y 999	Asp	gat Asp 970	att Ile	gtc Val	att Ile	tat Tyr	gat Asp 975	caa Gln	2928
acg Thr	gga Gly	Thr	aat Asn 980	gaa Glu	atc Ile	gag Glu	Arg	ttc Phe 985	cat His	ttt Phe	ccg Pro	cgt Arg	cag Gln 990	aat Asn	aag Lys	2976

gag Glu	cct Pro	tat Tyr 995	Leu	tgt Cys	ctt Leu	Ala	gat Asp 1000	Phe	ctt Leu	cgc Arg	Pro	gtt Val 1005	tcc Ser	agt Ser	G1y 999	3024
Glu	atg Met 1010	gac Asp	tat Tyr	gtt Val	Gly	ttc Phe 1015	ctt Leu	gct Ala	gta Val	Thr	gca Ala 1020	gga Gly	aaa Lys	Gly	att Ile	3072
cgt Arg 102	gaa Glu 5	tta Leu	999 999	gag Glu	cag Gln 103	Ala	aaa Lys	gag Glu	gct Ala	gga Gly 103	Asp	tat Tyr	tta Leu	ttc Phe	agt Ser 1040	3120
cac His	tta Leu	atc Ile	Gln	gca Ala 1045	aca Thr	gcc Ala	tta Leu	Glu	atg Met 1050	gcg Ala	gaa Glu	G1y 999	Phe	gcc Ala 1055	gag Glu	3168
cgt Arg	gtc Val	His	cag Gln LO60	ctc Leu	atg Met	cgt Arg	qaA	aag Lys 1065	tgg Trp	999 999	ttt Phe	Pro	gat Asp 1070	tcg Ser	gct Ala	3216
gac Asp	ttt Phe	aca Thr 075	atg Met	gaa Glu	gag Glu	Arg	ttc Phe 1080	gct Ala	gca Ala	aaa Lys	Tyr	cgt Arg 1085	ggc	atc Ile	cgt Arg	3264
Val	tcg Ser .090	ttt Phe	ggc Gly	tac Tyr	Pro	gca Ala 095	tgc Cys	cct Pro	gac Asp	Leu	gat Asp	gac Asp	caa Gln	gca Ala	aag Lys	3312
ttg Leu 1105	ttt Phe	aag Lys	ctg Leu	ttg Leu	aag Lys 1110	Pro	gga Gly	aag Lys	atc Ile	gga Gly 1115	Ile	gag Glu	ttg Leu	acg Thr	gaa Glu 1120	3360
61 Y 999	ttt Phe	atg Met	Met	gag Glu 125	cca Pro	gaa Glu	gcc Ala	Ser	gtc Val 130	acc Thr	gcg Ala	atg Met	Val	ttt Phe 135	gcc Ala	3408
cat His	cct Pro	Glu	gct Ala . 140	cgc Arg	tat Tyr	ttt Phe	Asn	gtt Val 145	tta Leu	tag						3441
<212	> 12 > 11 > PR' > Ba	T	us ha	alod	uran	8										

Met Thr Lys Ser Leu Phe Glu Gln Gln Leu Glu Arg Lys Ile Val Ile

Leu Asp Gly Ala Met Gly Thr Met Leu Gln Ala Ala Asn Leu Thr Ala

Asp Asp Phe Gly Gly Glu Glu Tyr Glu Gly Cys Asn Glu Tyr Leu Asn

Glu Thr Ala Pro His Val Val Glu Asp Ile His Arg Ala Tyr Leu Glu 55

Ala Gly Ala Asp Val Ile Ala Thr Asn Thr Phe Gly Ala Thr Asp Ile 65 75

- Val Leu Asp Asp Tyr Asp Leu Gly Tyr Lys Ala Glu Glu Leu Asn Ile 85 90 95
- Cys Ala Val Lys Ile Ala Lys Arg Val Ala Glu Glu Phe Ser Thr Pro 100 105 110
- Asp Trp Pro Arg Phe Val Ala Gly Ala Met Gly Pro Thr Thr Lys Ser
- Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Gln Leu Ile Glu Ser Tyr 130 135 140
- Arg Gln Gln Ala Thr Gly Leu Ile Lys Gly Gly Ala Asp Ile Leu Leu 145 150 155 160
- Leu Glu Thr Ser Gln Asp Met Arg Asn Val Lys Ala Ala Tyr Leu Gly
 165 170 175
- Leu Ser Gln Ala Gln Lys Glu Leu Glu Val Lys Leu Pro Leu Ile Ile 180 185 190
- Ser Gly Thr Ile Glu Pro Met Gly Thr Thr Leu Ala Gly Gln Asn Ile 195 200 205
- Glu Ala Phe Tyr Leu Ser Leu Glu His Met Asn Pro Val Val Val Gly 210 215 220
- Leu Asn Cys Ala Thr Gly Pro Glu Phe Met Arg Asp His Leu Arg Ser 225 230 235 240
- Leu Ser Asp Leu Ala Thr Cys Ser Val Ser Cys Tyr Pro Asn Ala Gly
 245 250 255
- Leu Pro Asp Glu Glu Gly Asn Tyr His Glu Ser Pro Glu Ser Leu Ala 260 265 270
- Ala Lys Leu Ala Gly Phe Ala Glu Lys Gly Trp Leu Asn Met Val Gly 275 280 285
- Gly Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Leu Leu Asp Val 290 295 300
- Met Lys Gln Phe Glu Pro Arg Gln Pro Lys Gly Asp His Pro His Ser 305 310 315 320
- Val Ser Gly Ile Glu Pro Leu Leu Tyr Asp Asp Ser Met Arg Pro Leu 325 330 335
- Phe Val Gly Glu Arg Thr Asn Val Ile Gly Ser Arg Lys Phe Lys Arg 340 345 350
- Leu Ile Glu Glu Glu Lys Tyr Glu Glu Ala Ser Glu Ile Ala Arg Ser 355 360 365
- Gln Val Lys Lys Gly Ala His Val Ile Asp Val Cys Leu Ala Asp Pro 370 375 380
- Asp Arg Asp Glu Met Glu Asp Met Glu Glu Phe Leu Lys Phe Val Ile 385 390 395 400
- Asn Lys Val Lys Val Pro Leu Met Ile Asp Ser Thr Asp Glu Lys Val

410

- Ile Glu Gln Ala Leu Thr Tyr Ser Gln Gly Lys Ala Ile Ile Asn Ser 420 425 430
- Ile Asn Leu Glu Asp Gly Glu Glu Arg Phe Glu Lys Val Val Pro Leu
 435 440 445
- Val His Lys Tyr Gly Ala Ala Val Val Val Gly Thr Ile Asp Glu Glu 450 460
- Gly Met Ala Ile Thr Ala Glu Lys Lys Leu Ala Val Ala Lys Arg Ser 465 470 475 480
- Tyr Asp Leu Leu Val Asn Lys Tyr Asn Ile Arg Pro Ser Asp Ile Ile 485 490
- Phe Asp Pro Leu Val Phe Pro Val Gly Thr Gly Asp Glu Gln Tyr Ile 500 505 510
- Gly Ser Ala Asn Glu Thr Val Glu Gly Ile Arg Arg Ile Lys Glu Glu 515 520 525
- Leu Pro Glu Cys Leu Thr Ile Leu Gly Val Ser Asn Val Ser Phe Gly 530 540
- Leu Pro Pro Val Gly Arg Glu Val Leu Asn Ala Ala Tyr Leu Tyr His 545 550 550 560
- Cys Thr Gln Ala Gly Leu Asp Tyr Ala Ile Val Asn Thr Glu Lys Leu 565 570 575
- Glu Arg Tyr Ala Ser Ile Ser Asp Glu Glu Lys Glu Leu Ser Arg Lys 580 585 590
- Leu Leu Phe Glu Thr Thr Asp Glu Thr Leu Ala Glu Phe Thr Ala Phe 595 600 605
- Tyr Arg Gly Lys Lys Ala Glu Lys Lys Val Glu Thr Ser Asn Leu Thr 610 615 620
- Leu Glu Glu Arg Leu Ala Asn Tyr Ile Val Glu Gly Ser Lys Asp Gly 625 630 640
- Leu Thr Glu Asp Leu Asp Lys Ala Leu Ala Lys Tyr Asp Asp Pro Leu 645 650 655
- Asp Ile Ile Asn Gly Pro Leu Met Asn Gly Met Asp Glu Val Gly Arg
 660 665 670
- Leu Phe Asn Asn Asn Glu Leu Ile Val Ala Glu Val Leu Gln Ser Ala 675 680 685
- Glu Val Met Lys Ala Ser Val Ala His Leu Glu Pro His Met Glu Lys 690 695 700
- Lys Ala Asp Asp His Gly Lys Gly Lys Ile Ile Leu Ala Thr Val Lys 705 710 715 720
- Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Glu Ile Ile Leu Ser 725 730 735

- Asn Asn Gly Phe Arg Ile Val Asn Leu Gly Ile Lys Val Thr Ser Asn 740 745 750
- Glu Leu Ile Glu Ala Val Ala Arg Glu Asn Pro Asp Ala Ile Gly Leu 755 760 765
- Ser Gly Leu Leu Val Lys Ser Ala Gln Gln Met Val Leu Thr Ala Gln 770 780
- Asp Leu Lys Gln Gln Gln Ile Ser Ile Pro Ile Leu Val Gly Gly Ala
 785 790 795 800
- Ala Leu Thr Arg Lys Phe Thr Asn Thr Lys Ile Ala Pro Glu Tyr Asp 805 810 815
- Gly Leu Val Val Tyr Ala Lys Asp Ala Met Asn Gly Leu Glu Leu Ala 820 825 830
- Asn Lys Leu Met Lys Pro Asp Glu Arg Glu Lys Leu Ala Val Ser Leu 835 840 845
- His Glu Ala Lys Glu Gln Ala Asn Ser Arg Thr Gln Met Gly Gly 850 855 860
- Gly Thr Ala Val Ala Val Lys Pro Thr Arg Ser His Val Ser Thr Thr 865 870 875 880
- Val Pro Val Ala Val Pro Pro Asp Val Lys Pro His Ile Leu Arg His 885 890 895
- His Ser Ile Ala His Leu Glu Pro Tyr Ile Asn Met Gln Met Leu Leu 900 905 910
- Gly Arg His Leu Gly Leu Gln Gly Lys Val Ser Arg Leu Leu Ala Glu 915 920 925
- Lys Asp Glu Lys Ala Leu Glu Leu Lys Glu Lys Val Asp Ala Leu Leu 930 935 940
- Thr Arg Val Lys Glu Glu Gln Leu Met Glu Ala His Gly Met Tyr Gln 945 950 955 960
- Phe Phe Pro Ala Gln Ser Asp Gly Asp Asp Ile Val Ile Tyr Asp Gln 965 970 975
- Thr Gly Thr Asn Glu Ile Glu Arg Phe His Phe Pro Arg Gln Asn Lys 980 985 990
- Glu Pro Tyr Leu Cys Leu Ala Asp Phe Leu Arg Pro Val Ser Ser Gly 995 1000 1005
- Glu Met Asp Tyr Val Gly Phe Leu Ala Val Thr Ala Gly Lys Gly Ile 1010 1015 1020
- Arg Glu Leu Gly Glu Gln Ala Lys Glu Ala Gly Asp Tyr Leu Phe Ser 1025 1030 1035 1040
- His Leu Ile Gln Ala Thr Ala Leu Glu Met Ala Glu Gly Phe Ala Glu
 1045 1050 1055
- Arg Val His Gln Leu Met Arg Asp Lys Trp Gly Phe Pro Asp Ser Ala 1060 1065 1070

384

432

110

										52						
ĄaĄ	Phe	Thr 1075	Met	: Glı	Glu	a Arg	Phe 1080		Ala	Lys	Tyr	Arg 1085		Ile	Arg	
Val	Ser 1090	Phe	Gly	туг	Pro	Ala 1095	Cys	Pro) Asp		Asp 1100		Gln	Ala	Lys	
Leu 110	Phe 5	Lys	Lev	Leu	Lys 1110	Pro	Gly	Lys		Gly 1115		Glu	Leu		Glu 1120	
Gly	Phe	Met	Met	Glu 1125	Pro	Glu	Ala		Val		Ala	Met		Phe 1135	Ala	
His	Pro		Ala 1140		Tyr	Phe	Asn	Val 1145								
<21 <21	0> 1 1> 3 2> D 3> B	411 NA	lus	stea	roth	ermo	phil	us								
<22																
	1> C 2> (DS 1)	(340	8)					•							
		BB02		-,												
	0> 1:	-														
atg Met 1	gct Ala	aac Asn	gtc Val	acc Thr 5	tta Leu	gaa Glu	cag Gln	caa Gln	ctg Leu 10	caa Gln	aga Arg	aaa Lys	att Ile	ctt Leu 15	gtc Val	48
atc Ile	gat Asp	ggc Gly	gcc Ala 20	atg Met	ggc Gly	acg Thr	atg Met	atc Ile 25	caa Gln	agc Ser	gcc Ala	aac Asn	cta Leu 30	tcg Ser	gcc Ala	96
gcc Ala	gac Asp	ttt Phe 35	ggc Gly	ggc Gly	gag Glu	gcg Ala	tat Tyr 40	gaa Glu	61Å 888	tgc Cys	aac Asn	gaa Glu 45	tat Tyr	ttg Leu	acc Thr	144
ctc Leu	acc Thr 50	gcc Ala	ccg Pro	cat His	gtc Val	atc Ile 55	cgc Arg	cgc Arg	att Ile	cat His	gaa Glu 60	gcg Ala	tac Tyr	cta Leu	gaa Glu	192
gcc Ala 65	ggt Gly	Ala	Asp	Ile	att Ile 70	Glu	acg Thr	Asn	acg Thr	Phe	gga Gly	gcg Ala	aca Thr	cgc Arg	atc Ile	240

gtg ctt gac gaa tat ggc ctc ggt cat ttg gcg ctt gag ctg aac atc

Val Leu Asp Glu Tyr Gly Leu Gly His Leu Ala Leu Glu Leu Asn Ile

gaa gcg gcc aaa ctc gcc aaa caa acg gct gag tcg ttc tcc acc ccg Glu Ala Ala Lys Leu Ala Lys Gln Thr Ala Glu Ser Phe Ser Thr Pro 105

gac tgg ccg cgc ttt gtc gcc ggt tcg atg ggg ccg acg acg aaa acg

Asp Trp Pro Arg Phe Val Ala Gly Ser Met Gly Pro Thr Thr Lys Thr 120

ttg tcg gtg aca ggc ggc gca acg ttt gaa gaa ctc gtc gcc gcc tac

Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Glu Leu Val Ala Ala Tyr

											140					
gaa Glu 145	Glu	caa Gln	gcg Ala	g cgc Arg	gga Gly 150	Leu	ctc Leu	tta Leu	gga Gly	ggc Gly 155	Val	gac Asp	ctt Leu	ctc Leu	cta Leu 160	480
cto Leu	gag Glu	acg Thr	tgc Cys	caa Gln 165	Asp	acg Thr	ctg Leu	aat Asn	gtc Val 170	Lys	gcc Ala	ggt Gly	ttt Phe	ctc Leu 175	ggc Gly	528
att Ile	tcg	aag Lys	gcg Ala 180	Phe	gaa Glu	gcg Ala	gtc Val	ggc Gly 185	cgc Arg	cgc Arg	gtg Val	ccg Pro	ctc Leu 190	atg Met	att Ile	576
tcc Ser	ggc Gly	acg Thr 195	atc Ile	gaa Glu	ccg Pro	atg Met	ggc Gly 200	acg Thr	acg Thr	ctc Leu	gcc Ala	999 Gly 205	cag Gln	gcg Ala	atc Ile	624
gat Asp	gcg Ala 210	ttt Phe	ttc Phe	atc Ile	tcg Ser	gtg Val 215	cgc Arg	cat His	atg Met	aag Lys	ccg Pro 220	atc Ile	gcc Ala	gtc Val	ggc Gly	672
Leu 225		Сув	Ala	Thr	Gly 230	Pro	Glu	Phe	Met	Thr 235	Asp	His	Leu	Arg	Thr 240	720
Leu	gcc Ala	Ser	Leu	Ala 245	Asp	Thr	Ala	Val	Ser 250	Сув	Tyr	Pro	Asn	Ala 255	Gly	768
Leu	ccg Pro	Asp	Glu 260	Glu	Gly	His	Tyr	His 265	Glu	Thr	Pro	Asn	Met 270	Leu	Ala	816
Glu	Lys	11e 275	Arg	Arg	Phe	Ala	Glu 280	Lys	Gly	Trp	Ile	Asn 285	Ile	Val	Gly	864
Gly	tgt Cys 290	Сув	Gly	Thr	Thr	Pro 295	Asp	His	Ile	Arg	Ala 300	Ile	Ala	Glu	Ala	912
Val 305	cgt Arg	Asp	Leu	Pro	Pro 310	Arg	Ala	Ile	Pro	Ser 315	Ser	Phe	qaA	Val	His 320	960
Ala	gtt Val	Ser	Gly	11e 325	Glu	Ala	Leu	Ile	Tyr 330	qsA	Glu	Thr	Met	Arg 335	Pro	1008
Leu	ttt Phe	Val	Gly 340	Glu	Arg	Thr	Asn	Val 345	Ile	Gly	Ser	Arg	Lys 350	Phe	Lys	1056
cgc Arg	ctc Leu	atc Ile 355	gcc Ala	gaa Glu	999 999	aaa Lys	tac Tyr 360	gaa Glu	gaa Glu	gcg Ala	gcg Ala	gaa Glu 365	atc Ile	gcc Ala	cgc Arg	1104
gcc Ala	caa Gln 370	gtg Val	aaa Lys	aac Asn	Gly	gcc Ala 375	cat His	gtc Val	atc Ile	gac Asp	att Ile 380	tgc Cys	ctc Leu	gcc Ala	gac Asp	1152

										54						
Pro 385) Ast	c cgo	g Asp	gaa Glu	Leu 390	His	gac Asp	atg Met	gag Glu	Gln 395	Phe	gtc Val	Arg	gaa Glu	gtc Val 400	1200
gtg Va]	l aaa Lys	aaa Lys	a gtg s Val	l aaa Lys 405	Val	ccg Pro	ctt Leu	gtc Val	Ile 410	Asp	tcg Ser	acc Thr	gac	gag Glu 415	Arg	1248
gto Val	atc Ile	gaa Glu	a cgc a Arg 420	Ala	ctt Leu	acg Thr	tat Tyr	Ser 425	Gln	999 999	aag Lys	gcg Ala	atc Ile 430	atc Ile	aac Asn	1296
t cg Ser	atc Ile	aac Asn 435	Leu	gaa Glu	gat Asp	ggc Gly	gaa Glu 440	Glu	cgg Arg	ttt Phe	gcg Ala	aag Lys 445	gtc Val	gtt Val	cct Pro	1344
ctc Leu	ctg Leu 450	His	caa Gln	tac	ggc Gly	gcc Ala 455	gcc Ala	gtt Val	gtc Val	gtc Val	ggc Gly 460	acg	atc Ile	gat Asp	gag Glu	1392
caa Gln 465	Gly	atg Met	gcg	gtt Val	aca Thr 470	gcc Ala	gaa Glu	cgg Arg	aaa Lys	ttg Leu 475	Glu	atc Ile	gcc Ala	ttg Leu	cgt Arg 480	1440
tcg Ser	tat Tyr	gac Asp	ttg Leu	ctg Leu 485	gtg Val	aac Asn	cgc Arg	tac Tyr	ggc Gly 490	gtc Val	ccc Pro	gag Glu	cgc Arg	gac Asp 495	atc Ile	1488
att Ile	ttc Phe	gac Asp	ccg Pro 500	ctc Leu	gtc Val	ttc Phe	ccg Pro	gtc Val 505	Gly	acc Thr	ggc	gat Asp	gag Glu 510	caa Gln	tac Tyr	1536
atc Ile	ggc Gly	gcg Ala 515	gcg Ala	aaa Lys	gaa Glu	acc Thr	att Ile 520	gag Glu	ggc Gly	atc Ile	cgc Arg	ctc Leu 525	att Ile	aaa Lys	gag Glu	1584
cgg Arg	ctg Leu 530	cct Pro	cat His	tgc Cys	ttg Leu	acg Thr 535	atg Met	ctt Leu	ggc Gly	atc Ile	agc Ser 540	aac Asn	gtc Val	tcg Ser	ttc Phe	1632
ggc Gly 545	ttg Leu	ccg Pro	ccg Pro	gcc Ala	gga Gly 550	cgc Arg	gag Glu	gtg Val	ctc Leu	aac Asn 555	tcc Ser	gtc Val	ttt Phe	ttg Leu	tac Tyr 560	1680
cat His	tgc Cys	acg Thr	caa Gln	gcc Ala 565	ggg Gly	ctc Leu	gat Asp	tac Tyr	gcc Ala 570	atc Ile	gtc Val	aac Asn	acc Thr	gag Glu 575	aaa Lys	1728
ttg Leu	gag Glu	cgg Arg	ttc Phe 580	gcc Ala	tcg Ser	att Ile	ccg Pro	gaa Glu 585	gag Glu	gaa Glu	gtg Val	cga Arg	atg Met 590	gct Ala	gag Glu	1776
gca Ala	ctt Leu	ctt Leu 595	ttt Phe	gac Asp	aca Thr	naA	gac Asp 600	gaa Glu	aca Thr	tta Leu	aac Asn	gcc Ala 605	ttt Phe	atc Ile	gaa Glu	1824
ttt Phe	tac Tyr 610	cga Arg	agc Ser	aaa Lys	Ile	acc Thr 615	gcc Ala	gcc Ala	aaa Lys	ccg Pro	gcg Ala 620	cag Gln	acg Thr	aac Asn	ttg Leu	1872
agc Ser	ttg Leu	gaa Glu	gag Glu	cgg Arg	ctc Leu	gcc Ala	cgc Arg	tac Tyr	gtt Val	att Ile	gaa Glu	gjå aaa	tcg Ser	aaa Lys	gac Asp	1920

	W	O 03/0	87386	5												PCT/EP03
										55		٠.		•		
625					630					635					640	•
Gly	, Let	ı Ile	: Lev	gat Asp 645	Leu	Glu	Lys	Ala	650	Glu	Thr	Tyr	Ser	Asp 655	Pro	1
ctg Leu	tcc Ser	atc Ile	Ile 660	aac Asn	ggt	ccg Pro	ctc Leu	atg Met 665	Ala	ggc Gly	atg Met	gat Asp	gaa Glu 670	Val	99 9 Gly	2016
cgg Arg	ctg Leu	Phe 675	Asn	aac Asn	aac Asn	cag Gln	ctc Leu 680	Ile	gtc Val	gct Ala	gaa Glu	gta Val 685	Leu	caa Gln	agc Ser	206 4
gcg Ala	gaa Glu 690	Val	atg Met	aaa Lys	gca Ala	gcg Ala 695	gtc Val	gcc Ala	ttt Phe	tta Leu	gag Glu 700	ctg Leu	tat Tyr	atg Met	gaa Glu	2112
aag Lys 705	aaa Lys	gaa Glu	gga Gly	agc Ser	aca Thr 710	aaa Lys	gga Gly	aaa Lys	gtc Val	att Ile 715	ctc Leu	gcc Ala	acc Thr	gtc Val	aaa Lys 720	2160
ggc Gly	gat Asp	gtg Val	cat His	gac Asp 725	atc Ile	ggc Gly	aaa Lys	aac Asn	ttg Leu 730	gtc Val	gac Asp	atc Ile	att Ile	tta Leu 735	agc Ser	2208
aac Asn	aac Asn	ggc	tac Tyr 740	gag Glu	gtg Val	atc Ile	gac Asp	ctc Leu 745	ggc Gly	att Ile	aaa Lys	gtc Val	gct Ala 750	Pro	cag Gln	2256
caa Gln	ctc Leu	att Ile 755	gaa Glu	gcg Ala	gtg Val	cgc Arg	gaa Glu 760	cat His	cag Gln	ccg Pro	gac Asp	atc Ile 765	atc Ile	G1y 999	ttg Leu	2304
tcg Ser	ggc Gly 770	ttg Leu	ctt Leu	gtg Val	aaa Lys	tcg Ser 775	gct Ala	caa Gln	cag Gln	atg Met	gtc Val 780	gtc Val	acc Thr	gcc Ala	caa Gln	2352
Авр 785	Leu	Arg	Gln	gcg Ala	Gly 790	Ile	Ser	Thr	Pro	Ile 795	Leu	Val	Gly	Gly	Ala 800	2400
Ala	Leu	Thr	Arg	aaa Lys 805	Phe	Thr	Glu	Asn	Lys 810	Ile	Ala	Pro	Glu	Tyr 815	Asp	2448
Gly	Val	Val	Leu 820	tac Tyr	Ala	Lys	Asp	Ala 825	Met	Asp	Gly	Leu	Ala 830	Leu	Ala	2496
Asn	Gln	Ile 835	Gln	cag Gln	Gly	Glu	Ile 840	Asp	Tyr	Lys	Lys	Lys 845	Glu	Thr	Ala	2544
Glu	agc Ser 850	gag Glu	cca Pro	acg Thr	cgg Arg	caa Gln 855	acg Thr	acg Thr	gtg Val	gtc Val	aca Thr 860	gcg Ala	gtc Val	aaa Lys	tcg Ser	2592
acc Thr 865	gtc Val	tcg Ser	acc Thr	gac Asp	gtt Val 870	ccc Pro	gtc Val	tac Tyr	atc Ile	ccg Pro 875	gcc Ala	gat Asp	ctc Leu	gag Glu	cgc Arg 880	2640

										56						
Hi	c gcg s Ala	g ctg Lev	ı Arg	aat Asr 889	ı Val	Pro	ctt Leu	gac Asp	His 890	: Ile	ttg Leu	ccg Pro	tac Tyr	gtc Val 895	aac Asn	2688
tgg	g caa o Glr	atg Met	gtg Val 900	Lev	ggc Gly	cac Kis	cac His	Leu 905	Gly	ttg Leu	aaa Lys	gga Gly	aaa Lys 910	Val	aaa Lys	2736
Arg	, Leu	915	Glu	Glu	aaa Lys	Asp	920	Lys	Ala	Leu	Ala	Leu 925	Lys	Ala	Val	2784
val	930	Glu	Leu	Leu	gcc Ala	Glu 935	Ala	Lys	Glu	Arg	Arg 940	Trp	Ile	Gln	Pro	2832
945	GIY	Val	Tyr	Arg	Phe 950	Phe	Pro	Ala	Gln	Ser 955	Asp	Gly	Asn	Arg	Val 960	2880
Tyr	lle	Tyr	Asp	965	act Thr	Asp	Gly	Lys	Thr 970	Val	Leu	Glu	Met	Phe 975	Asp	2928
Phe	Pro	Arg	980	Pro	cgg Arg	Ala	Pro	Tyr 985	Leu	Сув	Leu	Ala	990	Tyr	Leu	2976
гув	Ser	Lys 995	Glu	Ser	ggc	Glu ;	Met 1000	Asp	Tyr	Val	Gly	Leu 1005	Phe	Ala	Val	3024
Thr	Ala 1010	Gly	His	Gly		Arg .015	Glu	Leu	Ala	Gln	Arg 1020	Trp	Lys	Glu	Glu	3072
1025	gin Gin	Phe	Leu	Lys	agc Ser 1030	His	Ala	Ile	Gln	Ala 1035	Leu	Ala	Leu	Glu	Ile 1040	3120
AIA	Glu	Gly	Phe 1	Ala .045	gaa Glu	Arg	Ile	His 1	Gln .050	Ile	Met	Arg	Asp	Arg 1055	Trp	3168
GIÀ	Phe	Pro 1	Asp .060	Asp	ccg Pro	Asp	Phe 1	Thr .065	Met	Glu	Glu	Arg 1	Phe 070	Ala	Ala	3216
гув	Tyr 1	075	GIÀ	Gln	cgc Arg	Tyr 1	Ser 080	Phe	Gly	Tyr	Pro 1	Ala 1085	Сув	Pro	Asn	3264
Leu 1	090	Asp	Gln	Glu		Leu 095	Phe	Arg	Leu	Leu 1	His 100	Pro	Glu	qaA	Ile	3312
Gly 1105	Ile	Arg	Leu '	Thr	gac (Asp (1110	Gly '	Tyr	Met	Met	Glu 1115	Pro	Glu	Ala	Ser	Val 1120	3360
tcg Ser	gcg Ala	atc Ile	gtc Val	ttc Phe	gcc d Ala I	cat (ccg Pro	gaa Glu	gcg Ala	cgg Arg	tat Tyr	ttc Phe	aat Asn	gtg Val	tta Leu	3408

1125

1130

1135

taa

3411

<210> 14

<211> 1136

<212> PRT

<213> Bacillus stearothermophilus

<400> 14

Met Ala Asn Val Thr Leu Glu Gln Gln Leu Gln Arg Lys Ile Leu Val 1 5 10 15

Ile Asp Gly Ala Met Gly Thr Met Ile Gln Ser Ala Asn Leu Ser Ala 20 25 30

Ala Asp Phe Gly Gly Glu Ala Tyr Glu Gly Cys Asn Glu Tyr Leu Thr 35 40 45

Leu Thr Ala Pro His Val Ile Arg Arg Ile His Glu Ala Tyr Leu Glu 50 55

Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr Phe Gly Ala Thr Arg Ile
65 70 75 80

Val Leu Asp Glu Tyr Gly Leu Gly His Leu Ala Leu Glu Leu Asn Ile 85 90 95

Glu Ala Ala Lys Leu Ala Lys Gln Thr Ala Glu Ser Phe Ser Thr Pro 100 105 110

Asp Trp Pro Arg Phe Val Ala Gly Ser Met Gly Pro Thr Thr Lys Thr 115 120 125

Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Glu Leu Val Ala Ala Tyr 130 135 140

Glu Glu Gln Ala Arg Gly Leu Leu Leu Gly Gly Val Asp Leu Leu Leu 145 150 160

Leu Glu Thr Cys Gln Asp Thr Leu Asn Val Lys Ala Gly Phe Leu Gly 165 170 175

Ile Ser Lys Ala Phe Glu Ala Val Gly Arg Arg Val Pro Leu Met Ile 180 185 190

Ser Gly Thr Ile Glu Pro Met Gly Thr Thr Leu Ala Gly Gln Ala Ile 195 200 205

Asp Ala Phe Phe Ile Ser Val Arg His Met Lys Pro Ile Ala Val Gly 210 215 220

Leu Asn Cys Ala Thr Gly Pro Glu Phe Met Thr Asp His Leu Arg Thr 225 230 235 240

Leu Ala Ser Leu Ala Asp Thr Ala Val Ser Cys Tyr Pro Asn Ala Gly 245 250 255

Leu Pro Asp Glu Glu Gly His Tyr His Glu Thr Pro Asn Met Leu Ala 260 265 270 WO 03/087386 PCT/EP03/04010

58

Glu Lys Ile Arg Arg Phe Ala Glu Lys Gly Trp Ile Asn Ile Val Gly 275 280 285

- Gly Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Ile Ala Glu Ala 290 295 300
- Val Arg Asp Leu Pro Pro Arg Ala Ile Pro Ser Ser Phe Asp Val His 305 310 315 320
- Ala Val Ser Gly Ile Glu Ala Leu Ile Tyr Asp Glu Thr Met Arg Pro 325 330 335
- Leu Phe Val Gly Glu Arg Thr Asn Val Ile Gly Ser Arg Lys Phe Lys 340 345 350
- Arg Leu Ile Ala Glu Gly Lys Tyr Glu Glu Ala Ala Glu Ile Ala Arg 355 360 365
- Ala Gln Val Lys Asn Gly Ala His Val Ile Asp Ile Cys Leu Ala Asp 370 380
- Pro Asp Arg Asp Glu Leu His Asp Met Glu Gln Phe Val Arg Glu Val
 385 390 395 400
- Val Lys Lys Val Lys Val Pro Leu Val Ile Asp Ser Thr Asp Glu Arg
 405 410 415
- Val Ile Glu Arg Ala Leu Thr Tyr Ser Gln Gly Lys Ala Ile Ile Asn 420 425 430
- Ser Ile Asn Leu Glu Asp Gly Glu Glu Arg Phe Ala Lys Val Val Pro 435 440 445
- Leu Leu His Gln Tyr Gly Ala Ala Val Val Gly Thr Ile Asp Glu 450 455 460
- Gln Gly Met Ala Val Thr Ala Glu Arg Lys Leu Glu Ile Ala Leu Arg 465 470 475 480
- Ser Tyr Asp Leu Leu Val Asn Arg Tyr Gly Val Pro Glu Arg Asp Ile 485 490 495
- Ile Phe Asp Pro Leu Val Phe Pro Val Gly Thr Gly Asp Glu Gln Tyr 500 505 510
- Ile Gly Ala Ala Lys Glu Thr Ile Glu Gly Ile Arg Leu Ile Lys Glu
 515 520 525
- Arg Leu Pro His Cys Leu Thr Met Leu Gly Ile Ser Asn Val Ser Phe 530 540
- Gly Leu Pro Pro Ala Gly Arg Glu Val Leu Asn Ser Val Phe Leu Tyr 545 550 555 560
- His Cys Thr Gln Ala Gly Leu Asp Tyr Ala Ile Val Asn Thr Glu Lys
 565 570 575
- Leu Glu Arg Phe Ala Ser Ile Pro Glu Glu Glu Val Arg Met Ala Glu 580 585 590
- Ala Leu Leu Phe Asp Thr Asn Asp Glu Thr Leu Asn Ala Phe Ile Glu
 595 600 605

Phe	Tyr 610	Arg	Ser	Lys	Ile	Thr 615	Ala	Ala	Lys	Pro	Ala 620	Gln	Thr	Asn	Leu
Ser 625	Leu	Glu	Glu	Arg	Leu 630	Ala	Arg	Tyr	Val	Ile 635	Glu	Gly	Ser	Lys	Asp 640
Gly	Leu	Ile	Leu	Asp 645	Leu	Glu	Lys	Ala	Leu 650	Glu	Thr	Tyr	Ser	Авр 655	Pro
Leu	Ser	Ile	Ile 660	Asn	Gly	Pro	Leu	Met 665	Ala	Gly	Met	Asp	Glu 670	Val	Gly
Arg	Leu	Phe 675	Asn	Asn	Asn	Gln	Leu 680	Ile	Val	Ala	Glu	Val 685	Leu	Gln	Ser
Ala	Glu 690	Val	Met	Lys	Ala	Ala 695	Val	Ala	Phe	Leu	Glu 700	Leu	Tyr	Met	Glu
Lys 705	Lys	Glu	Gly	Ser	Thr 710	Lys	Gly	Lys	Val	Ile 715	Leu	Ala	Thr	Val	Lys 720
Gly	Asp	Val	His	Asp 725	Ile	Gly	Lys	Asn	Leu 730	Val	Авр	Ile	Ile	Leu 735	Ser
Asn	Asn	Gly	Tyr 740	Glu	Val	Ile	Asp	Leu 745	Gly	Ile	Lys	Val	Ala. 750	Pro	Gln
Gln	Leu	Ile 755	Glu	Ala	Val	Arg	Glu 760	His	Gln	Pro	Asp	Ile 765	Ile	Gly	Leu
Ser	Gly 770	Leu	Leu	Val	Lys	Ser 775	Ala	Gln	Gln	Met	Val 780	Val	Thr	Ala	Gln
Авр 785		Arg	Gln	Ala	Gly 790	Ile	Ser	Thr	Pro	Ile 795	Leu	Val	Gly	Gly	Ala 800
Ala	Leu	Thr	Arg	Lys 805	Phe	Thr	Glu	Asn	Lys 810	Ile	Ala	Pro	Glu	Tyr 815	Asp
Gly	Val	Val	Leu 820	Tyr	Ala	Lys	Ąsp	Ala 825	Met	Asp	Gly	Leu	Ala 830	Leu	Ala
Asn		Ile 835	Gln	Gln	Gly	Glu	Ile 840	Asp	Tyr	Lys	Lys	Lys 845	Glu	Thr	Ala
Glu	Ser 850	Glu	Pro	Thr	Arg	Gln 855	Thr	Thr	Val	Val	Thr 860	Ala	Val	ГÀв	Ser
Thr 865	Val	Ser	Thr	Asp	Val 870	Pro	Val	Tyr	Ile	Pro 875	Ala	Asp	Leu	Glu	Arg 880
His	Ala	Leu	Arg	Asn 885	Val	Pro	Leu	Asp	His 890	Ile	Leu	Pro	Tyr	Val 895	Asn
Trp	Gln	Met	Val 900	Leu	Gly	His	His	Leu 905	Gly	Leu	Lys	Gly	Lys 910	Val	Lys
Arg	Leu	Leu 915	Glu	Glu	Lys	Asp	Glu 920	Lys	Ala	Leu	Ala	Leu 925	Lys	Ala	Val

Val Asp Glu Leu Leu Ala Glu Ala Lys Glu Arg Arg Trp Ile Gln Pro

930

935

940

Ala Gly Val Tyr Arg Phe Phe Pro Ala Gln Ser Asp Gly Asn Arg Val 945 950 955 960

Tyr Ile Tyr Asp Pro Thr Asp Gly Lys Thr Val Leu Glu Met Phe Asp 965 970 975

Phe Pro Arg Gln Pro Arg Ala Pro Tyr Leu Cys Leu Ala Asp Tyr Leu 980 985 990

Lys Ser Lys Glu Ser Gly Glu Met Asp Tyr Val Gly Leu Phe Ala Val 995 1000 1005

Thr Ala Gly His Gly Val Arg Glu Leu Ala Gln Arg Trp Lys Glu Glu 1010 1015 1020

Gly Glu Phe Leu Lys Ser His Ala Ile Gln Ala Leu Ala Leu Glu Ile 1025 1030 1035 1040

Ala Glu Gly Phe Ala Glu Arg Ile His Gln Ile Met Arg Asp Arg Trp
1045 1050 1055

Gly Phe Pro Asp Asp Pro Asp Phe Thr Met Glu Glu Arg Phe Ala Ala 1060 1065 1070

Lys Tyr Gln Gly Gln Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn 1075 1080 1085

Leu Glu Asp Gln Glu Lys Leu Phe Arg Leu Leu His Pro Glu Asp Ile 1090 1095 1100

Gly Ile Arg Leu Thr Asp Gly Tyr Met Met Glu Pro Glu Ala Ser Val 1105 1110 1115 1120

Ser Ala Ile Val Phe Ala His Pro Glu Ala Arg Tyr Phe Asn Val Leu 1125 1130 1135

<210> 15

<211> 3681

<212> DNA

<213> Vibrio cholerae

<220>

<221> CDS

<222> (1)..(3678)

<223> RVC04265

<400> 15

gtg gga aaa gaa gta aga caa caa ctc gaa cag caa ttg aaa caa cgt 48
Val Gly Lys Glu Val Arg Gln Gln Leu Glu Gln Gln Leu Lys Gln Arg
1 5 10 15

atc cta ctg att gat ggt ggt atg ggt acc atg att cag agt tat aag 96
Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Lys
20 25 30

tta caa gag gaa gac tat cgc ggt gca cga ttt gtc gat tgg cac tgt 144

										61						
Let	ı Glı	a Gl:	u Glv 5	. Asj	Туг	Arg	Gly 40		Arg	, bpe	Val	. Авр 45		His	Сув	
gat	ttg Lev 50	ı Lye	a gga a Gly	a aat / Asr	aac Asn	gac Asp 55	Leu	tta Leu	gtg Val	ctt Leu	act Thr 60	Gln	ccg Pro	Caa Gln	att Ile	192
65	Lys	GIV	ı Ile	His	70	Ala	Tyr	Leu	Glu	Ala 75	Gly	Ala	Aap	Ile	Ctt Leu 80	240
gag Glu	Thr	aac Asn	acc Thr	Phe 85	Asn	tca Ser	acc	acg Thr	att Ile 90	Ala	atg Met	gca Ala	gac Asp	tat Tyr 95	Asp	288
atg Met	Caa Gln	tcg Ser	Ctc Leu 100	Ser	gct Ala	gaa Glu	att	aac Asn 105	ttt Phe	gcc Ala	gcg Ala	gct Ala	aag Lys 110	ctt Leu	gca Ala	336
cgt Arg	gaa Glu	gtc Val 115	Ala	gat Asp	gag Glu	tgg Trp	acg Thr 120	gct Ala	aaa Lys	gat Asp	cca Pro	agt Ser 125	cgg Arg	cca Pro	cgc Arg	384
tat Tyr	gtg Val 130	Ala	ggt Gly	gtg Val	ctt Leu	999 Gly 135	cca Pro	acc Thr	aac Asn	cgt Arg	act Thr 140	tgc Cys	tct Ser	att Ile	tcg Ser	432
Pro 145	Asp	Val	Asn	Asp	cca Pro 150	Gly	Phe	Arg	Asn	Val 155	Thr	Phe	Авр	Gly	Leu 160	480
Val	Glu	Ala	Tyr	Ser 165	gaa Glu	Ser	Thr	Arg	Ala 170	Leu	Ile	Lys	Gly	Gly 175	Ser	528
Asp	Leu	Ile	Leu 180	Ile	gaa Glu	Thr	Ile	Phe 185	Asp	Thr	Leu	Asn	Ala 190	Lys	Ala	576
сув	Ala	Phe 195	Ala	Val	gat Asp	Ser	Val 200	Phe	Glu	Glu	Leu	Gly 205	Ile	Ser	Leu	624
cct Pro	gtg Val 210	atg Met	att Ile	tcc Ser	ggc Gly	acg Thr 215	att Ile	acc Thr	gat Asp	gcc Ala	tct Ser 220	Gly 999	cga Arg	act Thr	ctg Leu	672
tca Ser 225	gga Gly	cag Gln	aca Thr	acg Thr	gaa Glu 230	gct Ala	ttc Phe	tac Tyr	aac Asn	gcc Ala 235	ttg Leu	cgt Arg	cat His	gta Val	cgg Arg 240	720
ccg Pro	att Ile	tcg Ser	ttt Phe	ggc Gly 245	ttg Leu	aac Asn	tgt Cys	Ala	tta Leu 250	ggt Gly	cct Pro	gat Asp	gag Glu	ctg Leu 255	cgc Arg	768
cag Gln	tac Tyr	gtg Val	gaa Glu 260	gag Glu	ctt Leu	tca Ser	Arg	att Ile 265	tca Ser	gaa Glu	tgc Cys	Tyr	gtt Val 270	tcc Ser	gcg Ala	816
cac His	cca Pro	aat Asn 275	gcc Ala	gga Gly	ctg Leu	Pro	aat Asn 280	gcg Ala	ttt Phe	ggt Gly	Glu	tac Tyr 285	gat Asp	ctc Leu	tct Ser	864

	~~-																
	Ala	: gag : Gl: 290	ı Glı	a atg u Met	g gca : Ala	gaa Glu	cat His 295	Ile	gcg	gaa Glu	tgg Trp	gca Ala 300	Gln	gct Ala	ggc	ttt Phe	912
	ttg Leu 305	Ası	tto Lei	g gto 1 Val	ggt Gly	ggt Gly 310	tgc Cys	tgt Cys	gga Gly	act Thr	aca Thr 315	Pro	gag Glu	cat His	atc Ile	gcc Ala 320	960
	gcc Ala	att Ile	gco Ala	aaa Lys	gcc Ala 325	Val	gag Glu	ggt Gly	gta Val	aaa Lys 330	Pro	agg Arg	gct Ala	ctg Leu	cca Pro 335	gat Asp	1008
	ctg Leu	aaa Lys	gta Val	gaa Glu 340	Сув	cgt Arg	ctc Leu	tcg Ser	ggt Gly 345	Leu	gag Glu	ccg Pro	ctc Leu	aat Aen 350	Ile	ggt Gly	1056
	cct Pro	gaa Glu	acc Thr 355	Leu	ttt Phe	gtt Val	aac Asn	gtg Val 360	ggc	gaa Glu	cgt Arg	act Thr	aac Asn 365	gtc Val	acc Thr	ggt Gly	1104
	tct Ser	gcg Ala 370	Arg	ttt Phe	aag Lys	cgt Arg	tta Leu 375	att Ile	aaa Lys	gaa Glu	gag Glu	caa Gln 380	tac Tyr	gac Asp	gaa Glu	gcg Ala	1152
	Leu 385	Авр	Val	Ala	Arg	Glu 390	caa Gln	Val	Glu	Asn	Gly 395	Ala	Gln	Ile	Ile	Авр 400	1200
	Ile	Asn	Met	Asp	Glu 405	Gly	atg Met	Leu	Asp	Ala 410	Glu	Ala	Сув	Met	Val 415	Arg	1248
	ttt Phe	ttg Leu	aat Asn	cta Leu 420	tgc Cys	gcc Ala	tct Ser	gaa Glu	cca Pro 425	gaa Glu	ata Ile	tcc Ser	aaa Lys	gtt Val 430	ccg Pro	gtg Val	1296
	atg Met	gtc Val	gac Asp 435	tcc Ser	tct Ser	aaa Lys	tgg Trp	gaa Glu 440	gtc Val	att Ile	gaa Glu	gcg Ala	ggt Gly 445	ctg Leu	aaa Lys	tgc Сув	1344
	Ile	Gln 450	Gly	Lys	Gly	Ile	gtc Val 455	Asn	Ser	Ile	Ser	Leu 460	Lys	Glu	Gly	Lys	1392
•	gag Glu 465	aag Lys	ttt Phe	att Ile	gcc Ala	caa Gln 470	gcc Ala	aaa Lys	ttg Leu	gtg Val	cgc Arg 475	cgc Arg	tac Tyr	ggt Gly	gcc Ala	gcg Ala 480	1440
,	gtg Val	att Ile	gtg Val	atg Met	gca Ala 485	ttt Phe	gac Asp	gaa Glu	gtg Val	ggc Gly 490	caa Gln	gcc Ala	gat Asp	acc Thr	cgt Arg 495	gag Glu	1488
2	egc Arg	aaa Lys	tta Leu	gag Glu 500	atc Ile	tgt Cys	cgt Arg	Arg	gct Ala 505	tac Tyr	cat His	att Ile	ttg Leu	gtc Val 510	gat Asp	gag Glu	1536
9	gtg /al	ggc Gly	ttc Phe 515	cca Pro	ccg Pro	gaa Glu	gat Asp	att Ile 520	att Ile	ttt Phe	gac Asp	ccg Pro	aac Asn 525	atc Ile	ttt Phe	gct Ala	1584
9	jtt	gcg	acc	gga	att	gat	gag	cac	aat	aac	tac	gca	ctg	gat	ttc	att	1632

									6	33						
Val	Ala 530	Thr	Gly	Ile	Asp	Glu 535	His	Asn			Ala 540	Leu	Asp	Phe	Ile	•
aat Asn 545	gca Ala	gtg Val	gcg Ala	gac Asp	att Ile 550	aag Lys	cgt Arg	ga g Glu	ctg Leu	ccg Pro 555	cat His	gcg Ala	atg. Met	att Ile	tct Ser 560	1680
			tct Ser					Ser						Tyr		1728
			atc Ile 580													1776
atg Met	gac Asp	atg Met 595	99 9 Gly	att Ile	gtc Val	aac Asn	gca Ala 600	61 y 999	cag Gln	ctt Leu	gaa Glu	atc Ile 605	tac Tyr	gat Asp	aac Asn	1824
			aaa Lys				Ala									1872
			ggc Gly													1920
			gtt Val													1968
			gtg Val 660													2016
	_		atc Ile	_		_		-	_	_	_	_		_	_	2064
aaa Lys	cca Pro 690	ctg. Leu	gaa Glu	gtg Val	att Ile	gaa Glu 695	Gly 999	ccg Pro	ctg Leu	atg Met	gat Asp 700	ggt Gly	atg Met	aac Asn	gtg Val	2112
			ttg Leu													2160
			cgt Arg													2208
			caa Gln 740													2256
gcg Ala			aaa Lys													2304
			cag Gln													2352

gtg Val 785	cct Pro	tgc Cys	gag Glu	cag Gln	atc Ile 790	ctc Leu	aaa Lys	gtc Val	gca Ala	cgc Arg 795	gag Glu	caa Gln	aat Asn	gtc Val	gat Asp 800	2400
atc Ile	atc Ile	ggt Gly	ctc Leu	tct Ser 805	999 Gly	ctt Leu	atc Ile	acg Thr	ccg Pro 810	tct Ser	ttg Leu	gat Asp	gag Glu	atg Met 815	gta Val	2448
cac His	gtg Val	gcg Ala	aaa Lys 820	gag Glu	atg Met	gag Glu	cga Arg	caa Gln 825	G1y 999	ttt Phe	gaa Glu	ctg Leu	cca Pro 830	ctt Leu	ttg Leu	2496
att Ile	ggt Gly	999 Gly 835	gca Ala	aca Thr	acg Thr	tct Ser	aaa Lys 840	gcg Ala	cat His	act Thr	gcg Ala	gtg Val 845	aag Lys	att Ile	gaa Glu	2544
cag Gln	aat Asn 850	tat Tyr	cat His	gcg Ala	cct Pro	gta Val 855	gtg Val	tac Tyr	gtg Val	aat Asn	aac Asn 860	gcg Ala	tcg Ser	cgc Arg	gcg Ala	2592
			tgc Cys													2640
			ttg Leu													2688
cgt Arg	aaa Lys	acg Thr	CCC Pro 900	aaa Lys	tcg Ser	cgc Arg	cca Pro	gtc Val 905	acg Thr	tta Leu	gag Glu	cag Gln	gca Ala 910	cgt Arg	gct Ala	2736
naA	Lys	Ala 915	gcg Ala	Leu	Авр	Trp	Ala 920	Asn	Tyr	Thr	Pro	Pro 925	Ala	Pro	Ala	2784
Lys	Pro 930	Gly	gtg Val	His	Val	Phe 935	Glu	Asn	Ile	Ala	Leu 940	Ala	Thr	Leu	Arg	2832
Pro 945	Tyr	Ile	gat Asp	Trp	Thr 950	Pro	Phe	Phe	Met	Thr 955	Trp	Ser	Leu	Met	Gly 960	2880
Lys	Tyr	Pro	gcc Ala	Ile 965	Leu	Glu	His	Glu	Glu 970	Val	Gly	Glu	Glu	Ala 975	Lys	2928
Arg	Leu	Phe	cat His 980	Asp	Ala	Asn	Ala	Leu 985	Leu	Asp	Lys	Val	Glu 990	Arg	Glu	2976
gga Gly	cta Leu	ctg Leu 995	aaa Lys	gcc Ala	agt Ser	Gly	atg Met .000	tgt Cys	gca Ala	ctg Leu	Phe	cca Pro 1005	gca Ala	gcc Ala	agc Ser	3024
Val	ggc Gly 010	gat Asp	gac Asp	att Ile	Glu	gtg Val 015	tac Tyr	agt Ser	gat Asp	Glu	tcg Ser 020	cgt Arg	acg Thr	caa Gln	gtc Val	3072
gcg	cat	gtg	ctg	tac	aac	ttg	cgt	cag	cag	act	gag	aaa	ccg	aaa	999	3120

		•	55		
Ala His Val Leu 1025	Tyr Asn Leu 1030	Arg Gln Gln	Thr Glu Lys 1035	Pro Lys Gly 104	
gcc aac tac tgt Ala Asn Tyr Cys					
cgt gat tgg att Arg Asp Trp Ile 1060	Gly Ala Phe	gca gta act Ala Val Thr 1065.	Gly Gly Ile	ggt gag cga Gly Glu Arg 1070	3216
gcc ttg gcc gat Ala Leu Ala Asp 1075	Ala Tyr Lys	gct cag ggt Ala Gln Gly 1080	gat gat tac Asp Asp Tyr 1085	aat gcg atc Asn Ala Ile	3264
atg atc caa gcg Met Ile Gln Ala 1090	gta gcc gat Val Ala Asp 1095	cgt ttg gcg Arg Leu Ala	gaa gcc ttt Glu Ala Phe 1100	gcg gaa tat Ala Glu Tyr	3312
ctg cat gaa aaa Leu His Glu Lys 1105	gtg cgt aaa Val Arg Lys 1110	gag att tgg Glu Ile Trp	ggt tat gcg Gly Tyr Ala 1115	agc gat gaa Ser Asp Glu 112	
aat ctc tcc aat Asn Leu Ser Asn	gat gac ctg Asp Asp Leu 1125	atc cgt gag Ile Arg Glu 1130	cgt tat cag Arg Tyr Gln	ggc att cga Gly Ile Arg 1135	3408
ccc gcg ccg ggg Pro Ala Pro Gly 1140	Tyr Pro Ala		His Thr Glu		
ttg tgg cag atg Leu Trp Gln Met 1155	Leu Asn Val	gaa gag acc Glu Glu Thr 1160	ata ggt atg Ile Gly Met 1165	tca ctg acc Ser Leu Thr	3504
aca agc tat gcg Thr Ser Tyr Ala 1170	atg tgg ccg Met Trp Pro 1175	ggc gct tcg Gly Ala Ser	gta tcc ggt Val Ser Gly 1180	tgg tat ttc Trp Tyr Phe	3552
tcg cat ccc gat Ser His Pro Asp 1185	tct cgc tat Ser Arg Tyr 1190	ttt gcg gta Phe Ala Val	gcg cag atc Ala Gln Ile 1195	caa cca gat Gln Pro Asp 120	
caa ctg cac agc Gln Leu His Ser	tac gct gag Tyr Ala Glu 1205	cgt aaa ggt Arg Lys Gly 1210	tgg cgt ttg Trp Arg Leu	gaa gaa gct Glu Glu Ala 1215	3648
gaa aag tgg cta Glu Lys Trp Leu 1220	Ala Pro Asn		taa		3681
<210> 16 <211> 1226 <212> PRT <213> Vibrio ch	olerae				
400- 16					

Val Gly Lys Glu Val Arg Gln Gln Leu Glu Gln Gln Leu Lys Gln Arg

Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Lys

10

5

<400> 16

- Leu Gln Glu Glu Asp Tyr Arg Gly Ala Arg Phe Val Asp Trp His Cys
 35 40 45
- Asp Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ile 50 55 60
- Ile Lys Glu Ile His Ser Ala Tyr Leu Glu Ala Gly Ala Asp Ile Leu 65 70 75 80
- Glu Thr Asn Thr Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Asp
 85 90 95
- Met Gln Ser Leu Ser Ala Glu Ile Asn Phe Ala Ala Ala Lys Leu Ala 100 105 110
- Arg Glu Val Ala Asp Glu Trp Thr Ala Lys Asp Pro Ser Arg Pro Arg 115 120 125
- Tyr Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Cys Ser Ile Ser 130 135 140
- Pro Asp Val Asn Asp Pro Gly Phe Arg Asn Val Thr Phe Asp Gly Leu 145 150 155 160
- Val Glu Ala Tyr Ser Glu Ser Thr Arg Ala Leu Ile Lys Gly Gly Ser 165 170 175
- Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala 180 185 190
- Cys Ala Phe Ala Val Asp Ser Val Phe Glu Glu Leu Gly Ile Ser Leu 195 200 205
- Pro Val Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu 210 220
- Ser Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ala Leu Arg His Val Arg 225 230 235 240
- Pro Ile Ser Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg 245 250 255
- Gln Tyr Val Glu Glu Leu Ser Arg Ile Ser Glu Cys Tyr Val Ser Ala 260 265 270
- His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Ser 275 280 285
- Ala Glu Glu Met Ala Glu His Ile Ala Glu Trp Ala Gln Ala Gly Phe 290 295 300
- Leu Asn Leu Val Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala 305 310 315 320
- Ala Ile Ala Lys Ala Val Glu Gly Val Lys Pro Arg Ala Leu Pro Asp 325 330 335
- Leu Lys Val Glu Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly 340 345 350

-	_
	7
n	•

- Pro Glu Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly 355 360 365
- Ser Ala Arg Phe Lys Arg Leu Ile Lys Glu Glu Gln Tyr Asp Glu Ala 370 375 380
- Leu Asp Val Ala Arg Glu Gln Val Glu Asn Gly Ala Gln Ile Ile Asp 385 395 400
- Ile Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Cys Met Val Arg
 405 410 415
- Phe Leu Asn Leu Cys Ala Ser Glu Pro Glu Ile Ser Lys Val Pro Val 420 425 430
- Met Val Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys Cys 435 440 445
- Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Lys
 450 460
- Glu Lys Phe Ile Ala Gln Ala Lys Leu Val Arg Arg Tyr Gly Ala Ala 465 470 475 480
- Val Ile Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Arg Glu 485 490 495
- Arg Lys Leu Glu Ile Cys Arg Arg Ala Tyr His Ile Leu Val Asp Glu 500 505 510
- Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala 515 520 525
- Val Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile 530 535 540
- Asn Ala Val Ala Asp Ile Lys Arg Glu Leu Pro His Ala Met Ile Ser 550 555 560
- Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn Tyr Val 565 570 575
- Arg Glu Ala Ile His Ala Val Phe Leu Tyr His Cys Phe Lys His Gly
 580 585 590
- Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr Asp Asn 595 600 605
- Val Pro Leu Lys Leu Arg Glu Ala Val Glu Asp Val Ile Leu Asn Arg 610 620
- Arg Ser Asp Gly Thr Glu Arg Leu Leu Glu Ile Ala Glu Ala Tyr Arg 625 630 635 640
- Glu Asn Ser Val Gly Lys Glu Glu Asp Ala Ser Ala Leu Glu Trp Arg 645 650 655
- Ala Trp Pro Val Ala Lys Arg Leu Glu His Ala Leu Val Lys Gly Ile 660 665 670
- Thr Glu Phe Ile Val Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ser 675 680 685

- Lys Pro Leu Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val 690 695 700
- Val Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val 705 710 715 720
- Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe
 725 730 735
- Ile Asn Ala Gln Lys Ser Gly Ser Thr Ser Asn Gly Lys Ile Leu Leu 740 745 750
- Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly
 755 760 765
- Val Val Leu Gln Cys Asn Asn Phe Glu Ile Ile Asp Leu Gly Val Met
 770 780
- Val Pro Cys Glu Gln Ile Leu Lys Val Ala Arg Glu Gln Asn Val Asp 785 790 795 800
- Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val 805 810 815
- His Val Ala Lys Glu Met Glu Arg Gln Gly Phe Glu Leu Pro Leu Leu 820 825 830
- Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu 835 840 845
- Gln Asn Tyr His Ala Pro Val Val Tyr Val Asn Asn Ala Ser Arg Ala 850 855 860
- Val Gly Val Cys Thr Ser Leu Leu Ser Asp Glu Gln Arg Pro Gly Phe 865 870 875 880
- Ile Glu Arg Leu Asp Leu Asp Tyr Glu Arg Thr Arg Asp Gln His Ala 885 890 . 895
- Arg Lys Thr Pro Lys Ser Arg Pro Val Thr Leu Glu Gln Ala Arg Ala 900 905 910
- Asn Lys Ala Ala Leu Asp Trp Ala Asn Tyr Thr Pro Pro Ala Pro Ala 915 920 925
- Lys Pro Gly Val His Val Phe Glu Asn Ile Ala Leu Ala Thr Leu Arg 930 935 940
- Pro Tyr Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Met Gly 945 950 955 960
- Lys Tyr Pro Ala Ile Leu Glu His Glu Glu Val Gly Glu Glu Ala Lys 965 970 975
- Arg Leu Phe His Asp Ala Asn Ala Leu Leu Asp Lys Val Glu Arg Glu 980 985 990
- Gly Leu Leu Lys Ala Ser Gly Met Cys Ala Leu Phe Pro Ala Ala Ser 995 1000 1005
- Val Gly Asp Asp Ile Glu Val Tyr Ser Asp Glu Ser Arg Thr Gln Val

1010

1015

1020

Ala His Val Leu Tyr Asn Leu Arg Gln Gln Thr Glu Lys Pro Lys Gly 1025 1030 1035 1040

Ala Asn Tyr Cys Leu Ser Asp Tyr Val Ala Pro Lys Glu Ser Gly Lys 1045 1050 1055

Arg Asp Trp Ile Gly Ala Phe Ala Val Thr Gly Gly Ile Gly Glu Arg 1060 1065 1070

Ala Leu Ala Asp Ala Tyr Lys Ala Gln Gly Asp Asp Tyr Asn Ala Ile 1075 1080 1085

Met Ile Gln Ala Val Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr 1090 1095 1100

Leu His Glu Lys Val Arg Lys Glu Ile Trp Gly Tyr Ala Ser Asp Glu 1105 1110 1115 1120

Asn Leu Ser Asn Asp Asp Leu Ile Arg Glu Arg Tyr Gln Gly Ile Arg 1125 1130 1135

Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Ala Thr 1140 1145 1150

Leu Trp Gln Met Leu Asn Val Glu Glu Thr Ile Gly Met Ser Leu Thr 1155 1160 1165

Thr Ser Tyr Ala Met Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe 1170 1180

Ser His Pro Asp Ser Arg Tyr Phe Ala Val Ala Gln Ile Gln Pro Asp 1185 1190 1195 1200

Gln Leu His Ser Tyr Ala Glu Arg Lys Gly Trp Arg Leu Glu Glu Ala 1205 1210 1215

Glu Lys Trp Leu Ala Pro Asn Leu Asp Ala 1220 1225

<210> 17

<211> 3822

<212> DNA

<213> Sinorhizobium meliloti

<220>

<221> CDS

<222> (1)..(3819)

<223> RSM07338

<400> 17

gtg agt aaa tcg ata att ctt tgt cgt ttt cag aac ggg aga tct ccc 48
Val Ser Lys Ser Ile Ile Leu Cys Arg Phe Gln Asn Gly Arg Ser Pro
1 5 10 15

atg tcc gcc gcc gcc ctc ttt gga aac gtc tcg ccc aag ccg gat 96
Met Ser Ala Ala Asp Ala Leu Phe Gly Asn Val Ser Pro Lys Pro Asp
20 25 30

ggt tcg gaa gtc ttt cgg cag ctc gcc cag gcg gcg gct gaa cgc atc 14

Gly	Ser	Glu 35		Phe	: Arg	Gln	Leu 40		Gln	/U Ala	Ala	Ala 45		Arg	Ile	
Cto Leu	ato Ile 50	Met	gat Asp	ggc Gly	gcc Ala	atg Met 55	Gly	acg Thr	gag Glu	atc Ile	cag Gln 60	cag Gln	ctc Leu	ggt Gly	ttc Phe	192
gtg Val 65	Glu	gat Asp	cac His	ttc Phe	cgc Arg 70	ggc	gag Glu	cgc Arg	ttc Phe	ggt Gly 75	ggc	tgc Cys	gcc Ala	tgc Cys	cat His 80	240
Gln	Gln	Gly	Asn	Asn 85	Asp	ctc Leu	Leu	Thr	Leu 90	Thr	Gln	Pro	Lys	Ala 95	Ile	288
Glu	Asp	Ile	His 100	Tyr	His	tac Tyr	Ala	Ile 105	Ala	Gly	Ala	Asp	Ile 110	Leu	Glu	336
Thr	Asn	Thr 115	Phe	Ser	Ser	acg Thr	Arg 120	Ile	Ala	Gln	Ala	Asp 125	Tyr	Gly	Met	384
Glu	Asp 130	Met	Val	Tyr	Asp	ctc Leu 135	Asn	Arg	Asp	Gly	Ala 140	Arg	Leu	Ala	Arg	432
Arg 145	Ala	Ala	Lys	Arg	Ala 150	gag Glu	Ala	Glu	Asp	Gly 155	Arg	Arg	Arg	Phe	Val 160	480
Ala	Gly	Ala	Leu	Gly 165	Pro	acc Thr	Asn	Arg	Thr 170	Ala	Ser	Ile	Ser	Pro 175	Asp	528
Val	Asn	Asn	Pro 180	Gly	Tyr	cga Arg	Ala	Val 185	Ser	Phe	Asp	qaA	Leu 190	Arg	Leu	576
Ala	Tyr	Ala 195	Glu	Gln	Val	cgg Arg	Gly 200	Leu	Ile	Авр	Gly	Gly 205	Ala	qaA	Ile	624
TIE	210	Ile	Glu	Thr	Ile	ttc Phe 215	Asp	Thr	Leu	Asn	Ala 220	Lys	Ala	Ala	Ile	672
Phe 225	Ala	Thr	Gln	Glu	Val 230	ttt Phe	Ala	Glu	Lys	Gly 235	Val	Arg	Leu	Pro	Val 240	720
Met	Ile	Ser	Gly	Thr 245	Ile	acc Thr	Asp	Leu	Ser 250	Gly	Arg	Thr	Leu	Ser 255	Gly	768
Gln	Thr	Pro	Thr 260	Ala	Phe	tgg Trp	Tyr	Ser 265	Val	Arg	His	Ala	Авр 270	Pro	Phe	816
acg Thr	Ile	999 Gly : 275	ctc Leu	aac Asn	tgc Cys	gcg Ala	ctc (Leu (280	ggc Gly	gca Ala	aat Asn	Ala	atg Met 285	cgc Arg	gcc Ala	cat His	864

ata Ile	9ac 290	Glu	g ctt Leu	tcg Ser	g gcg	gto Val 295	Ala	gac	acg Thr	ctc Leu	gtc Val 300	Сув	gcc Ala	tat Tyr	ccg	912
aat Asr 305) Ala	ggc	ctg Leu	rce Pro	aac Asn 310	Glu	ttc Phe	ggc	cgc	tat Tyr 315	gac	gaa Glu	agc Ser	ccc	gag Glu 320	960
cag Gln	, atg Met	gcg Ala	gcg Ala	Gln 325	\Val	gag Glu	ggc	ttc Phe	gcc Ala 330	Arg	gac Asp	ggt Gly	ctc Leu	gtc Val 335	aac Asn	1008
ato Ile	gtc Val	ggc	ggc Gly 340	Cys	tgc Cys	ggt Gly	tcc Ser	acg Thr 345	ccg Pro	gcc Ala	cat His	atc Ile	cgc Arg 350	Ala	att Ile	1056
Ala	Glu	Ala 355	Val	Ala	Lys	Тут	ccg Pro 360	Pro	Arg	Arg	Val	Pro 365	Glu	Ile	Asp	1104
Arg	cgc Arg 370	atg Met	cgg Arg	ctt Leu	tcc Ser	ggc Gly 375	ctc Leu	gaa Glu	ccc Pro	ttc Phe	acg Thr 380	ctt Leu	acc Thr	gac Asp	gag Glu	1152
11e 385	Pro	Phe	Val	Asn	Val 390	Gly	gaa Glu	Arg	Thr	Asn 395	Val	Thr	Gly	Ser	Ala 400	1200
Lys	Phe	Arg	rys	Leu 405	Ile	Thr	gcc Ala	Gly	Asp 410	Tyr	Ala	Ala	Ala	Leu 415	Asp	1248
Val	Ala	Arg	Asp 420	Gln	Val	Ala	aat Asn	Gly 425	Ala	Gln	Ile	Ile	Авр 430	Val	Asn	1296
met	Asp	435	GIA	Leu	Ile	Asp	tcg Ser 440	Lys	Gln	Val	Met	Val 445	Glu	Phe	Leu	1344
Asn	Leu 450	Val	Ala	Ser	Glu	Pro 455	gat Asp	Ile	Ala	Arg	Val 460	Pro	Val	Met	Ile	1392
465	Ser	Ser	ГÀв	Trp	Glu 470	Val	atc Ile	Glu	Ala	Gly 475	Leu	Lys	Сув	Val	Gln 480	1440
Gly	Lys	Ala	Leu	Val 485	Asn	Ser	atc Ile	Ser	Leu 490	Lys	Glu	Gly	Glu	Ala 495	Ala	1488
ttc Phe	ctg Leu	cac His	cat His 500	gcg Ala	cgc Arg	ctc Leu	gtg Val	cgc Arg 505	gcc Ala	tat Tyr	ggc Gly	gcc Ala	gcg Ala 510	gtc Val	gtg Val	1536
gtg Val	Met	gcg Ala 515	ttc Phe	gac Asp	gag Glu	aag Lys	ggc Gly 520	cag Gln	gcc Ala	gac Asp	Thr	aaa Lys 525	acc Thr	cgc Arg	aag Lys	1584
gtg	gaa	atc	tgc	cgg	cgg	gcc	tat	cgg	ctg	ctg	acg	gaa	gag	gtt	ggc	1632

										72						
Va]	530		е Сув	Arç	g Arg	Ala 535		Arg	Leu	Leu	Thr 540		Glu	Val	Gly	
tto Phe 545	Pro	Pro	gag Glu	gad Asp	atc Ile 550	Ile	ttc Phe	: Asp	Pro	aat Asn 555	Ile	ttc Phe	gcg Ala	gtc Val	gcg Ala 560	1680
acc Thr	ggc Gly	ato Ile	gag Glu	gag Glu 569	g cac His	aac Asn	aat Asn	tac Tyr	ggc Gly 570	Val	gac Asp	ttc Phe	atc Ile	gag Glu 575	gcg Ala	1728
acg Thr	cac His	gag Glu	Ile 580	Ile	gcg Ala	gca Ala	ctg Leu	ccg Pro 585	cat His	gtc Val	cac His	gtc Val	tcc Ser 590	ggc Gly	ggc Gly	1776
gtg Val	tcg Ser	aac Asn 595	Leu	tcc Ser	ttt Phe	tcc Ser	ttc Phe 600	Arg	ggc Gly	aac Asn	gag Glu	ccg Pro 605	gtg Val	cgc Arg	gag Glu	1824
gcg Ala	atg Met 610	cac His	gcc Ala	atc Ile	ttc Phe	ctt Leu 615	tat Tyr	cac His	gcg Ala	atc Ile	cag Gln 620	gcc Ala	ggc Gly	atg Met	gac Asp	1872
atg Met 625	Gly	atc Ile	gtc Val	aat Asn	gcc Ala 630	gga Gly	cag Gln	ctc Leu	gcc Ala	gtc Val 635	tat Tyr	gat Asp	gcg Ala	atc Ile	gac Asp 640	1920
ccg Pro	gaa Glu	ctg Leu	cgc Arg	gaa Glu 645	acc Thr	tgc Cys	gag Glu	gac Asp	gtg Val 650	gtg Val	ctc Leu	aac Asn	cgc Arg	cgg Arg 655	gcc Ala	1968
gat Asp	tcg Ser	acc Thr	gag Glu 660	cgc Arg	ctc Leu	ctg Leu	gag Glu	atc Ile 665	gcc Ala	gag Glu	cgc Arg	tat Tyr	cgc Arg 670	999 999	aag Lys	2016
ggc Gly	61y 999	agc Ser 675	cag Gln	ggc Gly	aag Lys	gag Glu	aag Lys 680	gac Asp	ctt Leu	gcc Ala	tgg Trp	cgc Arg 685	gaa Glu	tgg Trp	ccg Pro	2064
gtg Val	gag Glu 690	aag Lys	cgg Arg	ctc Leu	gaa Glu	cac His 695	gcg Ala	ctc Leu	gtc Val	aat Asn	gga Gly 700	att Ile	acc Thr	gaa Glu	ttt Phe	2112
atc Ile 705	gaa Glu	gcc Ala	gat Asp	acg Thr	gaa Glu 710	gag Glu	gcc Ala	cgg Arg	ctt Leu	gcc Ala 715	gcc Ala	gag Glu	cgg Arg	ccg Pro	ctg Leu 720	2160
cat His	gtc Val	atc Ile	gaa Glu	ggc Gly 725	ccg Pro	ctg Leu	atg Met	gcc Ala	999 Gly 730	atg Met	aac Asn	gtc Val	gtg Val	ggc Gly 735	gat Asp	2208
ctc Leu	ttc Phe	ggt Gly	tcc Ser 740	ggc Gly	aag Lys	atg Met	ttc Phe	ctg Leu 745	ccg Pro	cag Gln	gtg Val	gtc Val	aag Lys 750	tcc Ser	gcc Ala	2256
cgg Arg	Val	atg Met 755	aag Lys	cag Gln	gcc Ala	gtt Val	gcg Ala 760	gtg Val	ctg Leu	ctc Leu	ccc Pro	cat His 765	atg Met	gag Glu	gag Glu	2304
gag Glu	aag Lys 770	cgc Arg	gcc Ala	aat Asn	Gly ggc	ggc Gly 775	ggc Gly	gag Glu	gcg Ala	cgc Arg	gag Glu 780	agt Ser	gcc Ala	ggc Gly	aag Lys	2352

at: 11: 78:	e Lei	g atq 1 Mei	g gcg t Ala	acc Thr	gtc Val 790	Lys	ggc	gac Bac	gtg Val	rac His	Asp	ato	ggc	aag Lye	aac Asn 800	· 2400
ate Ile	gto Val	ggd l Gly	gto Val	gtg Val 805	Leu	gcc Ala	tgc Cys	aac Asn	aat Asn 810	Tyr	gag Glu	atc	atc Ile	gac Asp 815	ctc Leu	2448
gg(Gl ₎	gto Val	atg L Met	g gtg : Val 820	Pro	tcg Ser	gct Ala	aag Lys	atc Ile 825	_ctc _Leu	gaa Glu	gtg Val	gcg Ala	cgc Arg 830	gaa Glu	cag Gln	2496
aag Lys	gto Val	9ac Asp 835	atc Ile	gtc Val	ggt Gly	ctt Leu	tcc Ser 840	ggc	ctc Leu	atc	acg	ccg Pro 845	tcg Ser	ctg Leu	gac Asp	2544
gag	atg Met 850	Ala	cat His	gtc Val	gct Ala	tcc Ser 855	gag Glu	ctc Leu	gaa Glu	cgg Arg	gag Glu 860	ggc Gly	ttc Phe	gat Asp	gtc Val	2592
Pro 865	Leu	Leu	atc Ile	Gly	Gly 870	Ala	Thr	Thr	Ser	Arg 875	Val	His	Thr	Ala	Val 880	2640
Lys	Ile	Asn	ccg Pro	885	Tyr	Ser	Leu	Gly	Gln 890	Thr	Val	Tyr	Val	Thr 895	Авр	2688
Ala	Ser	Arg	gcg Ala 900	Val	Gly	Val	Val	Ser 905	Ser	Leu	Leu	Ser	Pro 910	Glu	Val	2736
Arg	Asp	Ser 915	tac Tyr	Lys	Lys	Thr	Val 920	Arg	Ala	Glu	Tyr	Leu 925	Lys	Val	Ala	2784
Asp	930	His	gcc Ala	Arg	Asn	Glu 935	Ala	Glu	Lys	Arg	Arg 940	Leu	Pro	Leu	Ser	2832
945	Ala	Arg	gcg Ala	Asn	Ala 950	Phe	Arg	Ile	Asp	Trp 955	yab	Ala	His	Gln	Pro 960	2880
Lys	Val	Pro	tcc Ser	Phe 965	Leu	Gly	Thr	Arg	Val 970	Phe	Glu	Gly	Trp	Авр 975	Leu	2928
Ala	Glu	Leu	gcc Ala 980	Arg	Tyr	Ile	Asp	Trp 985	Thr	Pro	Phe	Phe	Gln 990	Thr	Trp	2976
Glu	Leu	Lys 995	gjå aaa	Val	Phe	Pro 1	Lys 000	Ile	Leu	Asp	Asp 1	Glu 005	Arg	Gln	Gly	3024
Ala 1	Ala .010	Ala	cgc Arg	Gln :	Leu :	Phe 015	Glu	Asp	Ala	Gln 1	Ala .020	Met	Val	Glu	Lys	3072
atc	gtg	gcc	gag	gca	tgg 1	ttc	gcc	ccg	aag	gcc	gtg	atc	ggc	ttc	tgg	3120

WO 03/087386 PCT/EP03/04010

		Ala	Glu	Ala			Ala	Pro				Ile	Gly	Phe	Trp	
102					103	_				103				•	1040	
Pro	Ala	Ala	agc Ser	atg Met 1045	Gly	gac Asp	gac Asp	Val	cgc Arg 1050	ctg Leu	ttt Phe	gcc Ala	Asp	gag Glu 1055	gtg Val	3168
cgc	gaa Glu	Ala	gag Glu 1060	ctt Leu	gcc Ala	acc Thr	Phe	ttc Phe 1065	acg Thr	ctc Leu	cgc Arg	Gln	cag Gln 1070	atg Met	gtg Val	3216
aag Lys	Arg	gac Asp 1075	ggc	cgg	ccg Pro	Asn	gtc Val 1080	gcc Ala	ctt Leu	gcc Ala	Авр	ttc Phe 1085	gtc Val	gcc Ala	ccg Pro	3264
Ala	gcg Ala 1090	agc Ser	ggc Gly	aag Lys	Arg	gac Asp 1095	tat Tyr	gtc Val	ggc Gly	Gly	ttc Phe	gtg Val	gtg Val	acg Thr	gcc Ala	3312
ggc Gly 110	Ile	gag Glu	gaa Glu	gtg Val	gcg Ala 1110	Ile	gcc Ala	gaa Glu	cgc Arg	ttc Phe 1115	Glu	cgg Arg	gcg Ala	aac Asn	gac Asp 1120	3360
gat Asp	tat Tyr	tcc Ser	tcg Ser	atc Ile 1125	atg Met	gtc Val	aag Lys	Ala	ctt Leu 130	gcg Ala	gac Asp	cgc Arg	Phe	gca Ala 1135	gag Glu	3408
gcc Ala	ttt Phe	Ala	gag Glu 140	cgc Arg	atg Met	cat His	Glu	tat Tyr l145	gtc Val	cgc Arg	aag Lys	Glu	ctc Leu 150	tgg Trp	ggc	3456
tat Tyr	Ala	ccg Pro 155	gac Asp	gaa Glu	gcc Ala	Phe	acg Thr 1160	ccg Pro	cag Gln	gaa Glu	Leu	atc Ile 165	gcc Ala	gag Glu	ccc Pro	3504
Tyr	gcc Ala 1170	ggc Gly	atc Ile	cgc Arg	Pro	gcg Ala 175	ccc Pro	ggc Gly	tac Tyr	Pro	gcg Ala 180	cag Gln	ccc Pro	gac Asp	cac His	3552
acg Thr 1189	Glu	aag Lys	gag Glu	acg Thr	ctt Leu 1190	Phe	cgg Arg	ctc Leu	ctg Leu	gat Asp 1195	Ala	gaa Glu	gcc Ala	gct Ala	atc Ile 1200	3600
ggc Gly	gtc Val	cgg Arg	ctc Leu 1	acc Thr 205	gag Glu	agc Ser	tat Tyr	Ala	atg Met 210	tgg Trp	ccg Pro	ggc Gly	Ser	tcg Ser 215	gta Val	3648
tcg Ser	ggc Gly	Leu	tat Tyr 220	gtc Val	ggc Gly	cac His	Pro	gat Asp 225	tcc Ser	tat Tyr	tac Tyr	Phe	ggc Gly 230	gtc Val	gca Ala	3696
aag Lys	Ile	gag Glu 235	cgc Arg	gat Asp	cag Gln	Val	gag Glu 240	gac Asp	tat Tyr	gcc Ala	Asp	cgc Arg 245	aag Lys	cgc Arg	atg Met	3744
Ser	gtc Val 250	cgc Arg	gag Glu	gtc Val	Glu	cgc Arg 255	tgg Trp	ctt Leu	tcg Ser	Pro	atc Ile 260	ctc Leu	aat Asn	tac Tyr	gtg Val	3792
ccg Pro 1265	Met	ccg Pro	gag Glu '	Thr	gaa Glu 1270	gcg Ala	gcg Ala	gag Glu	tag							3822

- <210> 18
- <211> 1273
- <212> PRT
- <213> Sinorhizobium meliloti
- <400> 18
- Val Ser Lys Ser Ile Ile Leu Cys Arg Phe Gln Asn Gly Arg Ser Pro 1 5 10 15
- Met Ser Ala Ala Asp Ala Leu Phe Gly Asn Val Ser Pro Lys Pro Asp
 20 25 30
- Gly Ser Glu Val Phe Arg Gln Leu Ala Gln Ala Ala Ala Glu Arg Ile 35 40 45
- Leu Ile Met Asp Gly Ala Met Gly Thr Glu Ile Gln Gln Leu Gly Phe 50 60
- Val Glu Asp His Phe Arg Gly Glu Arg Phe Gly Gly Cys Ala Cys His 65 70 75 80
- Gln Gln Gly Asn Asn Asp Leu Leu Thr Leu Thr Gln Pro Lys Ala Ile 85 90
- Glu Asp Ile His Tyr His Tyr Ala Ile Ala Gly Ala Asp Ile Leu Glu 100 105 110
- Thr Asn Thr Phe Ser Ser Thr Arg Ile Ala Gln Ala Asp Tyr Gly Met
 115 120 125
- Glu Asp Met Val Tyr Asp Leu Asn Arg Asp Gly Ala Arg Leu Ala Arg 130 135 140
- Arg Ala Ala Lys Arg Ala Glu Ala Glu Asp Gly Arg Arg Arg Phe Val
- Ala Gly Ala Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp 165 170 175
- Val Asn Asn Pro Gly Tyr Arg Ala Val Ser Phe Asp Asp Leu Arg Leu 180 185 190
- Ala Tyr Ala Glu Gln Val Arg Gly Leu Ile Asp Gly Gly Ala Asp Ile 195 200 205
- Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile 210 215 220
- Phe Ala Thr Gln Glu Val Phe Ala Glu Lys Gly Val Arg Leu Pro Val 225 235 240
- Met Ile Ser Gly Thr Ile Thr Asp Leu Ser Gly Arg Thr Leu Ser Gly 245 250 255
- Gln Thr Pro Thr Ala Phe Trp Tyr Ser Val Arg His Ala Asp Pro Phe 260 265 270
- Thr Ile Gly Leu Asn Cys Ala Leu Gly Ala Asn Ala Met Arg Ala His 275 280 285

- Ile Asp Glu Leu Ser Ala Val Ala Asp Thr Leu Val Cys Ala Tyr Pro 290 295 300
- Asn Ala Gly Leu Pro Asn Glu Phe Gly Arg Tyr Asp Glu Ser Pro Glu 305 310 315 320
- Gln Met Ala Ala Gln Val Glu Gly Phe Ala Arg Asp Gly Leu Val Asn 325 330 335
- Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Ala His Ile Arg Ala Ile 340 345 350
- Ala Glu Ala Val Ala Lys Tyr Pro Pro Arg Arg Val Pro Glu Ile Asp 355 360 365
- Arg Arg Met Arg Leu Ser Gly Leu Glu Pro Phe Thr Leu Thr Asp Glu 370 380
- Ile Pro Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala 385 390 395 400
- Lys Phe Arg Lys Leu Ile Thr Ala Gly Asp Tyr Ala Ala Ala Leu Asp 405 410 415
- Val Ala Arg Asp Gln Val Ala Asn Gly Ala Gln Ile Ile Asp Val Asn 420 425 430
- Met Asp Glu Gly Leu Ile Asp Ser Lys Gln Val Met Val Glu Phe Leu 435 440 445
- Asn Leu Val Ala Ser Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile 450 460
- Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys Cys Val Gln 465 470 475 480
- Gly Lys Ala Leu Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Ala Ala 485 490 495
- Phe Leu His His Ala Arg Leu Val Arg Ala Tyr Gly Ala Ala Val Val 500 510
- Val Met Ala Phe Asp Glu Lys Gly Gln Ala Asp Thr Lys Thr Arg Lys 515 520 525
- Val Glu Ile Cys Arg Arg Ala Tyr Arg Leu Leu Thr Glu Glu Val Gly 530 540
- Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala 545 550 555 560
- Thr Gly Ile Glu Glu His Asn Asn Tyr Gly Val Asp Phe Ile Glu Ala 565 570 575
- Thr His Glu Ile Ile Ala Ala Leu Pro His Val His Val Ser Gly Gly
 580 585 590
- Val Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu 595 600 605
- Ala Met His Ala Ile Phe Leu Tyr His Ala Ile Gln Ala Gly Met Asp 610 615 620

- Met Gly Ile Val Asn Ala Gly Gln Leu Ala Val Tyr Asp Ala Ile Asp 625 630 635 640
- Pro Glu Leu Arg Glu Thr Cys Glu Asp Val Val Leu Asn Arg Arg Ala 645 650 655
- Asp Ser Thr Glu Arg Leu Leu Glu Ile Ala Glu Arg Tyr Arg Gly Lys
- Gly Gly Ser Gln Gly Lys Glu Lys Asp Leu Ala Trp Arg Glu Trp Pro 675 680 685
- Val Glu Lys Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Glu Phe 690 695 700
- Ile Glu Ala Asp Thr Glu Glu Ala Arg Leu Ala Ala Glu Arg Pro Leu 705 710 715 720
- His Val Ile Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp
 725 730 735
- Leu Phe Gly Ser Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala
 740 745 750
- Arg Val Met Lys Gln Ala Val Ala Val Leu Leu Pro His Met Glu Glu 755 760 765
- Glu Lys Arg Ala Asn Gly Gly Gly Glu Ala Arg Glu Ser Ala Gly Lys
 770 780
- Ile Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn 795 800
- Ile Val Gly Val Val Leu Ala Cys Asn Asn Tyr Glu Ile Ile Asp Leu 805 810 815
- Gly Val Met Val Pro Ser Ala Lys Ile Leu Glu Val Ala Arg Glu Gln 820 825 830
- Lys Val Asp Ile Val Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp 835 840 845
- Glu Met Ala His Val Ala Ser Glu Leu Glu Arg Glu Gly Phe Asp Val 850 855 860
- Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val 865 870 875 880
- Lys Ile Asn Pro Arg Tyr Ser Leu Gly Gln Thr Val Tyr Val Thr Asp 885 890 895
- Ala Ser Arg Ala Val Gly Val Val Ser Ser Leu Leu Ser Pro Glu Val 900 905 910
- Arg Asp Ser Tyr Lys Lys Thr Val Arg Ala Glu Tyr Leu Lys Val Ala 915 920 925
- Asp Ala His Ala Arg Asn Glu Ala Glu Lys Arg Arg Leu Pro Leu Ser 930 935 940
- Gln Ala Arg Ala Asn Ala Phe Arg Ile Asp Trp Asp Ala His Gln Pro

945

950

9

Lys Val Pro Ser Phe Leu Gly Thr Arg Val Phe Glu Gly Trp Asp Leu 965 970 975

Ala Glu Leu Ala Arg Tyr Ile Asp Trp Thr Pro Phe Phe Gln Thr Trp 980 985 990

Glu Leu Lys Gly Val Phe Pro Lys Ile Leu Asp Asp Glu Arg Gln Gly
995 1000 1005

Ala Ala Ala Arg Gln Leu Phe Glu Asp Ala Gln Ala Met Val Glu Lys 1010 1015 1020

Ile Val Ala Glu Ala Trp Phe Ala Pro Lys Ala Val Ile Gly Phe Trp 1025 1030 1035 1040

Pro Ala Ala Ser Met Gly Asp Asp Val Arg Leu Phe Ala Asp Glu Val 1045 1050 1055

Arg Glu Ala Glu Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Met Val

Lys Arg Asp Gly Arg Pro Asn Val Ala Leu Ala Asp Phe Val Ala Pro 1075 1080 1085

Ala Ala Ser Gly Lys Arg Asp Tyr Val Gly Gly Phe Val Val Thr Ala 1090 1095 1100

Gly Ile Glu Glu Val Ala Ile Ala Glu Arg Phe Glu Arg Ala Asn Asp 1105 1110 1115 1120

Asp Tyr Ser Ser Ile Met Val Lys Ala Leu Ala Asp Arg Phe Ala Glu 1125 1130 1135

Ala Phe Ala Glu Arg Met His Glu Tyr Val Arg Lys Glu Leu Trp Gly
1140 1145 1150

Tyr Ala Pro Asp Glu Ala Phe Thr Pro Gln Glu Leu Ile Ala Glu Pro 1155 1160 1165

Tyr Ala Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His 1170 1175 1180

Thr Glu Lys Glu Thr Leu Phe Arg Leu Leu Asp Ala Glu Ala Ala Ile 1185 1190 1195 1200

Gly Val Arg Leu Thr Glu Ser Tyr Ala Met Trp Pro Gly Ser Ser Val 1205 1210 1215

Ser Gly Leu Tyr Val Gly His Pro Asp Ser Tyr Tyr Phe Gly Val Ala 1220 1225 1230

Lys Ile Glu Arg Asp Gln Val Glu Asp Tyr Ala Asp Arg Lys Arg Met 1235 1240 1245

Ser Val Arg Glu Val Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Val 1250 1255 1260

Pro Met Pro Glu Thr Glu Ala Ala Glu 1265 1270

<21 <21	0 > 1 1 > 3 2 > D 3 > E	684 NA	rich	ia c	oli									.		•
<22	1> C 2> ((368 905	1)				,					,, ^b	,		
	0> 1 age		aaa	gtg	gaa	caa	cta	cat	aca	Can	tta	aat		cat	200	40
Val 1	Ser	Ser	Lys	Val 5	Glu	Gln	Leu	Arg	Ala 10	Gln	Leu	Asn	Glu	Arg 15	Ile	48
Leu	Val	Leu	Asp 20		Gly	Met	Gly	Thr 25	Met	Ile	Gln	Ser	Tyr 30	Arg	Leu	96
aac Asn	gaa Glu	gcc Ala 35	Asp	ttt Phe	cgt Arg	ggt	gaa Glu 40	cgc	ttt Phe	gcc	gac Asp	tgg Trp 45	cca Pro	tgc Cys	gac Asp	144
ctc Leu	aaa Lys 50	ggc Gly	aac Asn	aac Asn	gac Asp	ctg Leu 55	ctg Leu	gta Val	ctc Leu	agt Ser	aaa Lys 60	ccg Pro	gaa Glu	gtg Val	atc Ile	192
gcc Ala 65	gct Ala	atc Ile	cac His	aac Asn	gcc Ala 70	Tyr	ttt Phe	gaa Glu	gcg Ala	ggc Gly 75	Ala	gat Asp	atc Ile	atc Ile	gaa Glu 80	240
acc Thr	aac Asn	acc Thr	ttc Phe	aac Asn 85	tcc Ser	acg Thr	acc Thr	att Ile	gcg Ala 90	atg Met	gcg Ala	gat Asp	tac Tyr	cag Gln 95	atg Met	288
gaa Glu	tcc Ser	ctg Leu	tcg Ser 100	gcg Ala	gaa Glu	atc Ile	aac Asn	ttt Phe 105	gcg Ala	gcg Ala	gcg Ala	aaa Lys	ctg Leu 110	gcg Ala	cga Arg	336
gct Ala	tgt Cys	gct Ala 115	gac Asp	gag Glu	tgg Trp	acc Thr	gcg Ala 120	cgc Arg	acg Thr	cca Pro	gag Glu	aaa Lys 125	ccg Pro	cgc Arg	tac Tyr	384
gtt Val	gcc Ala 130	ggt Gly	gtt Val	ctc Leu	ggc Gly	ccg Pro 135	acc Thr	aac Asn	cgc Arg	acg Thr	gcg Ala 140	tct Ser	att Ile	tct Ser	ccg Pro	432
gac Asp 145	gtc Val	aac Asn	gat Asp	ccg Pro	gca Ala 150	ttt Phe	cgt Arg	aat Asn	atc Ile	act Thr 155	ttt Phe	gac Asp	G1y 999	ctg Leu	gtg Val 160	480
gcg Ala	gct Ala	tat Tyr	cga Arg	gag Glu 165	tcc Ser	acc Thr	aaa Lys	gcg Ala	ctg Leu 170	gtg Val	gaa Glu	ggt Gly	ggc Gly	gcg Ala 175	gat Asp	528
Leu	Ile	Leu	Ile 180	gaa Glu	Thr	Val	Phe	Asp 185	Thr	Leu	Asn	Ala	Lys 190	Ala	Ala	576
gta Val	ttt Phe	gcg Ala	gtg Val	aaa Lys	acg Thr	gag Glu	ttt Phe	gaa Glu	gcg Ala	ctg Leu	ggc Gly	gtt Val	gag Glu	ctg Leu	ccg Pro	624

		195	•				200)				205		,		
att Ile	atg Met 210	Ile	tco Ser	ggc Gly	acc Thr	atc Ile 215	Thr	gac	gcc Ala	tcc Ser	999 Gly 220	Arg	acg Thr	ctc	tcc Ser	672
999 Gly 225	Gln	acc Thr	Thr	gaa Glu	gca Ala 230	Phe	tac Tyr	aac Asn	tca Ser	Leu 235	Arg	cac His	gcc Ala	gaa Glu	gct Ala 240	720
ctg Leu	acc	ttt Phe	ggc	Ctg Leu 245	Asn	tgt Cys	gcg Ala	ctg Leu	999 Gly 250	ccc Pro	gat Asp	gaa Glu	ctg Leu	cgc Arg 255	cag Gln	768
tac Tyr	gtg Val	cag Gln	gag Glu 260	Leu	tca Ser	cgg Arg	att Ile	gcg Ala 265	gaa Glu	tgc Cys	tac Tyr	gtc Val	acc Thr 270	gcg Ala	cac His	816
ccg Pro	aac Asn	gcc Ala 275	61 999	cta Leu	ccc Pro	aac Asn	gcc Ala 280	ttt	ggt Gly	gag Glu	tac Tyr	gat Asp 285	ctc Leu	gac	gcc Ala	864
Asp	Thr 290	Met	Ala	Lys	Gln	11e 295	Arg	gaa Glu	Trp	Ala	Gln 300	Ala	Gly	Phe	Leu	912
Asn 305	Ile	Val	Gly	Gly	Cys 310	Сув	Gly	acc Thr	Thr	Pro 315	Gln	His	Ile	Āla	Ala 320	960
Met	Ser	Arg	Ala	Val 325	Glu	Gly	Leu	gcg Ala	Pro 330	Arg	Lys	Leu	Pro	Glu 335	Ile	1008
Pro	Val	Ala	Сув 340	Arg	Leu	Ser	Gly	ctg Leu 345	Glu	Pro	Leu	Asn	Ile 350	Gly	Glu	1056
Asp	Ser	Leu 355	Phe	Val	Asn	Val	Gly 360	gaa Glu	Arg	Thr	Asn	Val 365	Thr	Gly	Ser	1104
Ala	Lув 370	Phe	Lys	Arg	Leu	11e 375	Lys	gaa Glu	Glu	Lys	Tyr 380	Ser	Glu	Ala	Leu	1152
gat Asp 385	Val	Ala	Arg	Gln	Gln 390	Val	Glu	Asn	Gly	Ala 395	Gln	Ile	Ile	Asp	Ile 400	1200
aac Asn	Met	Asp	Glu	Gly 405	Met	Leu	Asp	Ala	Glu 410	Ala	Ala	Met	Val	Arg 415	Phe	1248
ctc Leu	Asn	Leu	11e 420	Ala	Gly	Glu	Pro	Asp 425	Ile	Ala	Arg	Val	Pro 430	Ile	Met	1296
atc (gac Asp	tcc Ser 435	tca Ser	aaa Lys	tgg Trp	qaA	gtc Val 440	att Ile	gaa Glu	aaa Lys	ggt Gly	ctg Leu 445	aag Lys	tgt Cys	atc Ile	1344

										81						
Gli	999 Gly 450	r Lys	ggc Gly	att Ile	gtt Val	aad Asr 455	sez	ato Ile	tcg Ser	atg Met	aaa Lys 460	Glu	ggc	gtc Val	gat Asp	1392
gco Ala 465	Phe	ato : Ile	cat His	cac His	909 Ala 470	Lys	ttg Leu	ttg Leu	cgt Arg	cgc Arg 475	Tyr	ggt	gcg	gca Ala	gtg Val 480	1440
gtg Val	gta Val	atg Met	gcc Ala	Phe 485	: Asp	gaa Glu	cag Gln	gga Gly	Cag Gln 490	Ala	gat Asp	act Thr	cgc	gca Ala 495	cgg	1488
aaa Lys	atc Ile	gag Glu	att Ile 500	Cys	cgt Arg	cgg Arg	gcg Ala	tac Tyr 505	Lys	atc Ile	ctc Leu	acc Thr	gaa Glu 510	Glu	gtt Val	1536
Gly	Phe	ccg Pro 515	Pro	Glu	Asp	Ile	1le 520	Phe	Asp	Pro	Asn	11e 525	Phe	Ala	Val	1584
Ala	Thr 530	ggc Gly	Ile	Glu	Glu	His 535	Asn	Asn	Tyr	Ala	Gln 540	Авр	Phe	Ile	Gly	1632
Ala 545	Сув	gaa Glu	qaA	Ile	Lys 550	Arg	Glu	Leu	Pro	His 555	Ala	Leu	Ile	Ser	Gly 560	1680
Gly	Val	tct Ser	Asn	Val 565	Ser	Phe	Ser	Phe	Arg 570	Gly	Asn	Авр	Pro	Val 575	Arg	1728
Glu	Ala	att Ile	His 580	Ala	Val	Phe	Leu	Tyr 585	Tyr	Ala	Ile	Arg	Asn 590	Gly	Met	1776
Asp	Met	999 Gly 595	Ile	Val	Asn	Ala	Gly 600	Gln	Leu	Ala	Ile	Tyr 605	Asp	Авр	Leu	1824
Pro	Ala 610	gaa Glu	Leu	Arg	Asp	Ala 615	Val	Glu	Asp	Val	1le 620	Leu	Asn	Arg	Arg	1872
Asp 625	Asp	ggc Gly	Thr	Glu	Arg 630	Leu	Leu	Glu	Leu	Ala 635	Glu	Lys	Tyr	Arg	Gly 640	1920
Ser	Lys	acc Thr	Asp	Asp 645	Thr	Ala	Asn	Ala	Gln 650	Gln	Ala	Glu	Trp	Arg 655	Ser	1968
tgg Trp	gaa Glu	gtg Val	aat Asn 660	aaa Lys	cgt Arg	ctg Leu	gaa Glu	tac Tyr 665	tcg Ser	ctg Leu	gtc Val	aaa Lys	ggc Gly 670	att Ile	acc Thr	2016
gag Glu	ttt Phe	atc Ile 675	gag Glu	cag Gln	gat Asp	acc Thr	gaa Glu 680	gaa Glu	gcc Ala	cgc Arg	cag Gln	cag Gln 685	gct Ala	acg Thr	cgc Arg	2064
ccg Pro	att Ile	gaa Glu	gtg Val	att Ile	gaa Glu	ggc Gly	ccg Pro	ttg Leu	atg Met	gac Asp	ggc Gly	atg Met	aat Asn	gtg Val	gtc Val	2112

	050	•				093)				700	*				
999 Gl ₃ 705	/ ysb	cto Lei	ttt Phe	ggc Gly	gaa Glu 710	Gly	aaa Lys	atg Met	tto Phe	ctg Leu 715	Pro	cag Gln	gtg Val	gtc Val	aaa Lys 720	2160
Ser	g gcg Ala	Arg	gto Val	atg Met 725	Lys	Cag Gln	gcg Ala	gtg Val	gcc Ala 730	Tyr	ctc Leu	gaa Glu	ccg Pro	ttt Phe 735	att Ile	2208
gaa Glu	gcc Ala	ago Ser	Lys 740	Glu	cag Gln	ggc	aaa Lys	acc Thr 745	Asn	ggc	aag Lys	atg Met	gtg Val 750	atc Ile	gcc Ala	2256
acc	gtg Val	aag Lys 755	Gly	gac Asp	gtc Val	cac His	gac Asp 760	Ile	ggt Gly	aaa Lys	aat Asn	atc Ile 765	gtt Val	ggt Gly	gtg Val	2304
Val	Ctg Leu 770	Gln	Сув	Asn	Asn	Туг 775	Glu	Ile	Val	Asp	Leu 780	Gly	Val	Met	Val	2352
Pro 785		Glu	Lys	Ile	Leu 790	Arg	Thr	Ala	Lys	Glu 795	Val	Asn	Ala	Asp	Leu 800	2400
Ile	ggc Gly	Leu	Ser	Gly 805	Leu	Ile	Thr	Pro	Ser 810	Leu	yab	Glu	Met	Val 815	Asn	2448
Val	gcg Ala	Lys	Glu 820	Met	Glu	Arg	Gln	Gly 825	Phe	Thr	Ile	Pro	Leu 830	Leu	Ile	2496
GIÀ	ggc	835	Thr	Thr	Ser	Lys	Ala 840	His	Thr	Ala	Val	Lys 845	Ile	Glu	Gln	2544
Asn	tac Tyr 850	Ser	Gly	Pro	Thr	Val 855	Tyr	Val	Gln	neA	Ala 860	Ser	Arg	Thr	Val	2592
Gly 865	gtg Val	Val	Ala	Ala	Leu 870	Leu	Ser	Asp	Thr	Gln 875	Arg	Asp	Двр	Phe	Val 880	2640
Ala	cgt Arg	Thr	Arg	Lys 885	Glu	Tyr	Glu	Thr	Val 890	Arg	Ile	Gln	His	Gly 895	Arg	2688
Lys	aaa Lys	Pro	Arg 900	Thr	Pro	Pro	Val	Thr 905	Leu	Glu	Ala	Ala	Arg 910	Asp	Asn	2736
Asp	ttc Phe	Ala 915	Phe	Asp	Trp	Gln	Ala 920	Tyr	Thr	Pro	Pro	Val 925	Ala	His	Arg	2784
ctc Leu	ggc Gly 930	gtg Val	cag Gln	gaa Glu	Val	gaa Glu 935	gcc Ala	agc Ser	atc Ile	Glu	acg Thr 940	ctg Leu	cgt Arg	aat Asn	tac Tyr	2832

										83				•		
atc Ile 945	Asp	tgg Trp	aca Thr	ccg Pro	ttc Phe 950	ttt Phe	atg Met	acc Thr	tgg Trp	teg Ser 955	ctg Leu	gcc Ala	99 9 Gly	aag . Lys	tat Tyr 960	2880
ccg Pro	cgc Arg	att Ile	ctg Leu	gaa Glu 965	gat Asp	gaa Glu	gtg Val	gtg Val	ggc Gly 970	gtt Val	gag Glu	gcg Ala	cag Gln	cgg Arg 975	ctg Leu	2928
ttt Phe	aaa Lys	gac Asp	gcc Ala 980	aac Asn	gac Asp	atg Met	ctg Leu	gat Asp 985	aaa Lys	tta Leu	agc Ser	gcc Ala	gag Glu 990	aaa Lys	acg Thr	2976
ctg Leu	aat Asn	ccg Pro 995	cgt Arg	ggc Gly	gtg Val	Val	ggc Gly 1000	ctg Leu	ttc Phe	ccg Pro	Ala	aac Asn 1005	cgt Arg	gtg Val	ggc	3024
Asp	gac Asp 1010	att Ile	gaa Glu	atc Ile	tac Tyr	cgt Arg 1015	gac Asp	gaa Glu	acg Thr	Arg	acc Thr 1020	cat His	gtg Val	atc Ile	aac Asn	3072
gtc Val 1025	Ser	cac His	cat His	ctg Leu	cgt Arg 1030	Gln	cag Gln	acc Thr	gaa Glu	aaa Lys 1035	Thr	ggc Gly	ttc Phe	gct Ala	aac Asn 1040	3120
Tyr	Сув	Leu	Ala 1	Авр .045	ttc Phe	Val	Ala	Pro	Lys 1050	Leu	Ser	Gly	Lys	Ala 1055	Asp	3168
Tyr	Ile	Gly 1	Ala .060	Phe	gcc Ala	Val	Thr	Gly .065	Gly	Leu	Glu	Glu 1	Asp 1070	Ala	Leu	3216
Ala	Asp	Ala 1075	Phe	Glu	gcg Ala	Gln 1	His 080	Asp	Asp	Tyr	Asn 1	Lys 1085	Ile	Met	Val	3264
Lys 1	Ala .090	Leu	Ala	Asp		Leu 095	Ala	Glu	Ala	Phe 1	Ala 100	Glu	Tyr	Leu	His	3312
Glu 1105	Arg	Val	Arg	Lys	gtc Val 1110	Tyr	Trp	Gly	Tyr	Ala 1115	Pro	Asn	Glu	Asn	Leu 1120	3360
Ser	Asn	Glu	Glu 1	Leu 125	atc Ile	Arg	Glu	Asn 1	Tyr 130	Glņ	Gly	Ile	Arg	Pro 1135	Ala	3408
ccg Pro	ggc Gly	Tyr	Pro 1	gcc Ala	tgc Cys	ccg Pro	Glu	cat His 145	acg Thr	gaa Glu	aaa Lys	Ala	acc Thr 150	atc Ile	tgg Trp	3456
gag Glu	Leu 1	Leu 155	Glu '	Val	Glu	Lys 1	His 160	Thr	Gly	Met	Lys 1	Leu 165	Thr	Glu	Ser	3504
Phe	gcc Ala 170	atg ' Met '	tgg (Trp 1	Pro		gca Ala 175	tcg Ser	gtt Val	tcg Ser	Gly	tgg Trp 180	tac Tyr	ttc Phe	agc Ser	cac His	3552
ccg Pro	gac Asp	agc : Ser :	aag (Lys :	tac Iyr '	tac :	gct Ala	gta Val	gca Ala	caa Gln	att Ile	cag Gln	cgc Arg	gat Asp	cag Gln	gtt Val	3600

1185

1190

1195

1200

gaa gat tat gcc cgc cgt aaa ggt atg agc gtt acc gaa gtt gag cgc 3648 Glu Asp Tyr Ala Arg Arg Lys Gly Met Ser Val Thr Glu Val Glu Arg 1205 1210 1215

tgg ctg gca ccg aat ctg ggg tat gac gcg gac tga Trp Leu Ala Pro Asn Leu Gly Tyr Asp Ala Asp 1220 1225

3684

<210> 20 <211> 1227 <212> PRT

<213> Escherichia coli

<400> 20

Val Ser Ser Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile 1 5 10 15

Leu Val Leu Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu 20 25 30

Asn Glu Ala Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp
45

Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile
50 60

Ala Ala Ile His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu 65 70 75 80

Thr Asn Thr Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Gln Met 85 90 95

Glu Ser Leu Ser Ala Glu Ile Asn Phe Ala Ala Ala Lys Leu Ala Arg 100 105 110

Ala Cys Ala Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Tyr
115 120 125

Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro 130 135 140

Asp Val Asn Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gly Leu Val

Ala Ala Tyr Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp 165 170 175

Leu Ile Leu Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala 180 185 190

Val Phe Ala Val Lys Thr Glu Phe Glu Ala Leu Gly Val Glu Leu Pro 195 200 205

Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser 210 220

Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala 225 230 235 240 WO 03/087386 PCT/EP03/04010

85

Leu Thr Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln 245 250 255

Tyr Val Gln Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His 260 265 270

Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala 275 280 285

Asp Thr Met Ala Lys Gln Ile Arg Glu Trp Ala Gln Ala Gly Phe Leu 290 295 300

Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Gln His Ile Ala Ala 305 310 315 320

Met Ser Arg Ala Val Glu Gly Leu Ala Pro Arg Lys Leu Pro Glu Ile 325 330 335

Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Glu 340 345 350

Asp Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser 355 360 365

Ala Lys Phe Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu 370 375 380

Asp Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile 385 390 395 400

Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe
405 410 415

Leu Asn Leu Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met 420 425 430

Ile Asp Ser Ser Lys Trp Asp Val Ile Glu Lys Gly Leu Lys Cys Ile 435

Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Asp 450 455 460

Ala Phe Ile His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val 465 470 475 480

Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Ala Arg
485 490 495

Lys Ile Glu Ile Cys Arg Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val

Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val 515 520 525

Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly 530 540

Ala Cys Glu Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly
550 555 560

Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg 565 570 575

- Glu Ala Ile His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met 580 590
- Asp Met Gly Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asp Leu 595 600 605
- Pro Ala Glu Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg 610 615 620
- Asp Asp Gly Thr Glu Arg Leu Leu Glu Leu Ala Glu Lys Tyr Arg Gly 625 635 640
- Ser Lys Thr Asp Asp Thr Ala Asn Ala Gln Gln Ala Glu Trp Arg Ser 645 650 655
- Trp Glu Val Asn Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr
 660 665 670
- Glu Phe Ile Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Thr Arg 675 680 685
- Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val 690 695 700
- Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys 705 710 715 720
- Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile 725 730 735
- Glu Ala Ser Lys Glu Gln Gly Lys Thr Asn Gly Lys Met Val Ile Ala 740 745 750
- Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val 755 760 765
- Val Leu Gln Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val 770 780
- Pro Ala Glu Lys Ile Leu Arg Thr Ala Lys Glu Val Asn Ala Asp Leu 785 790 795 800
- Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn 805 810 815
- Val Ala Lys Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile 820 825 830
- Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln 835 840 845
- Asn Tyr Ser Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val 850 855 860
- Gly Val Val Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val 865 870 875 880
- Ala Arg Thr Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Gly Arg 885 890 895
- Lys Lys Pro Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn

900

910

Asp Phe Ala Phe Asp Trp Gln Ala Tyr Thr Pro Pro Val Ala His Arg 915 920 925

- Leu Gly Val Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr 930 935 940
- Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr 955 960
- Pro Arg Ile Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu 965 970 975
- Phe Lys Asp Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Thr 980 985 990
- Leu Asn Pro Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly 995 1000 1005
- Asp Asp Ile Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Ile Asn 1010 1015 1020
- Val Ser His His Leu Arg Gln Gln Thr Glu Lys Thr Gly Phe Ala Asn 1025 1030 1035 1040
- Tyr Cys Leu Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp 1045 1050 1055
- Tyr Ile Gly Ala Phe Ala Val Thr Gly Gly Leu Glu Glu Asp Ala Leu 1060 1065 1070
- Ala Asp Ala Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val 1075 1080 1085
- Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His 1090 1095 1100
- Glu Arg Val Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Asn Leu 1105 1110 1115 1120
- Ser Asn Glu Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala 1125 1130 1135
- Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Ala Thr Ile Trp 1140 1145 1150
- Glu Leu Leu Glu Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser 1155 1160 1165
- Phe Ala Met Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His 1170 1175 1180
- Pro Asp Ser Lys Tyr Tyr Ala Val Ala Gln Ile Gln Arg Asp Gln Val 1185 1190 1195 1200
- Glu Asp Tyr Ala Arg Arg Lys Gly Met Ser Val Thr Glu Val Glu Arg 1205 1210 1215
- Trp Leu Ala Pro Asn Leu Gly Tyr Asp Ala Asp 1220 1225

<21 <21 <21 <22 <22 <22 <22 <22	0> 1> C 2> (3> R	771 NA almo DS 1)	(376		phim	uriu	m									
atg	Ser	cat	gtt Val	gcc Ala 5	cgt Arg	tgt Cys	tct Ser	ctt Leu	ttc Phe 10	cgc Arg	cag Gln	cac His	gct Ala	ttg Leu 15	tgc Cys	48
cag Gln	tat Tyr	ggc Gly	tcg Ser 20	tta Leu	cgt Arg	gga Gly	gcg Ala	ttg Leu 25	tcg Ser	gga Gly	gcg Ala	agt Ser	gtg Val 30	agc Ser	agc Ser	96
aaa Lys	gtt Val	gaa Glu 35	caa Gln	ctg Leu	cgt Arg	gcg Ala	cag Gln 40	tta Leu	aat Asn	gaa Glu	cgt Arg	att Ile 45	ctg Leu	gtg Val	ctg Leu	144
						atg Met 55										192
Asp 65	Phe	Arg	Gly	Glu	Arg 70	ttt Phe	Ala	Asp	Trp	Pro 75	Сув	Asp	Leu	Lys	Gly 80	240
Asn	Asn	Asp	Leu	Leu 85	Val	ctc Leu	Ser	Lys	Pro 90	Glu	Val	Ile	Ala	Ala 95	Ile	288
His	Asn	Ala	Tyr 100	Phe	Glu	gct Ala	Gly	Ala 105	Asp	Ile	Ile	Glu	Thr 110	Aen	Thr	336
Phe	Asn	Ser 115	Thr	Thr	Ile	gcg Ala	Met 120	Ala	ysb	Tyr	Arg	Met 125	Glu	Ser	Leu	384
Ser	Ala 130	Glu	Ile	Asn	Tyr	gcg Ala 135	Ala	Ala	Lys	Leu	Ala 140	Arg	Ala	Сув	Ala	432
Asp 145	Glu	Trp	Thr	Ala	Arg 150	aca Thr	Pro	Glu	Lys	Pro 155	Arg	Phe	Val	Ala	Gly 160	480
Val	Leu	Gly	Pro	Thr 165	Asn	cgc Arg	Thr	Ala	Ser 170	Ile	Ser	Pro	Asp	Val 175	Asn	528
Asp	Pro	Ala	Phe 180	Arg	Asn	atc Ile	Thr	Phe 185	Asp	Gln	Leu	Val	Ala 190	Ala	Tyr	576
cgt Arg	gaa Glu	tcc Ser	acc Thr	aaa Lys	gcg Ala	ctg Leu	gtg Val	gaa Glu	ggt Gly	ggc Gly	gca Ala	gat Asp	ctg Leu	att Ile	ctg Leu	624

195 200 205

		195	•				200)				205	i			
att Ile	gag Glu 210	Thr	gtt Val	ttt Phe	gac Asp	acc Thr 215	Lev	aat Asn	gcg Ala	, aaa Lys	gcg Ala 220	Ala	gtg Val	ttt Phe	gcg	672
gtg Val 225	Lys	gaa Glu	gag Glu	ttt Phe	gaa Glu 230	ı Ala	ctg Lev	ggc Gly	gtt Val	gac Asp 235	Leu	ccg Pro	ato Ile	atg Met	att Ile 240	720
tcc Ser	ggc Gly	acc Thr	atc Ile	acc Thr 245	Asp	gcc Ala	tct Ser	ggc	Arg 250	Thr	ctt Leu	tcc	ggc	Cag Gln 255	act Thr	768
acc Thr	gaa Glu	gcc	ttt Phe 260	Tyr	aac Asn	tcg Ser	Ctg	cgc Arg 265	His	gcc Ala	gag	gcg Ala	ctc Leu 270	Thr	ttt Phe	816
GIÀ	Leu	Asn 275	Сув	Ala	Leu		Pro 280	Asp	Glu	Leu	Arg	Gln 285	Tyr	Val	Gln	864
Glu	Leu 290	Ser	Arg	Ile	Ala	gaa Glu 295	Сув	Tyr	Val	Thr	Ala 300	His	Pro	Asn	Ala	912
Gly 305	Leu	Pro	Asn	Ala	Phe 310	ggc Gly	Glu	Tyr	Asp	Leu 315	Asp	Ala	Asp	Thr	Met 320	960
Ala	Lys	Gln	Ile	Arg 325	Glu	tgg Trp	Ala	Glu	Ala 330	Gly	Phe	Leu	Asn	11e 335	Val	1008
GIÀ	GIÀ	Сув	Сув 340	Gly	Thr	acg Thr	Pro	Glu 345	His	Ile	Ala	Ala	Met 350	Ser	Arg	1056
ATS	Val	355	Gly	Leu	Leu	Pro	Arg 360	Gln	Leu	Pro	Asp	11e 365	Pro	Val		1104
Сув	370	Leu	Ser	Gly	Leu	gag Glu 375	Pro	Leu	Asn	Ile	Gly 380	Asp	Asp	Ser	Leu	1152
Phe 385	Val	Asn	Val	Gly	Glu 390	cgt Arg	Thr	Asn	Val	Thr 395	Gly	Ser	Ala	Lys	Phe 400	1200
Lув	Arg	Leu	Ile	Lys 405	Glu	gag Glu	Lys	Tyr	Ser 410	Glu	Ala	Leu	Авр	Val 415	Ala	1248
Arg	Gln	Gln	Val 420	Glu	Ser	Gly ggc	Ala	Gln 425	Ile	Ile	Asp	Ile	Asn 430	Met	Asp	1296
gag Glu	GlY	atg Met 435	ctc Leu	gac Asp	gcc Ala	gaa Glu	gcg Ala 440	gcg Ala	atg Met	gtg Val	cgt Arg	ttc Phe 445	ctc Leu	agc Ser	ctg Leu	1344

									_	30						
att Ile	gcc Ala 450	ggt Gly	gag Glu	ccg Pro	gac Asp	att Ile 455	gcc Ala	cgt Arg	gta Val	cca Pro	atc Ile 460	atg Met	atc Ile	gac Asp	tcc Ser	1392
tcc Ser 465	aaa Lys	tgg Trp	gag Glu	gtt Val	atc Ile 470	gaa Glu	aaa Lys	999 999	ctg Leu	aag Lys 475	tgc Cya	att Ile	cag Gln	ggt Gly	aaa Lys 480	1440
ggc Gly	atc Ile	gtc Val	aac Asn	tct Ser 485	att Ile	tcg Ser	atg Met	aaa Lys	gag Glu 490	ggc Gly	gtg Val	gaa Glu	gcc Ala	ttt Phe 495	atț Ile	1488
cat His	cat His	gcg Ala	aag Lys 500	ctt Leu	ctg Leu	cgt Arg	cgc Arg	tac Tyr 505	ggc Gly	gcg Ala	gca Ala	gtg Val	gtg Val 510	gtg Val	atg Met	1536
gcc Ala	ttt Phe	gat Asp 515	gag Glu	cag Gln	G] A 333	cag Gln	gcc Ala 520	gat Asp	acc Thr	cgc Arg	gcg Ala	cgt Arg 525	aaa Lys	atc Ile	gaa Glu	1584
att Ile	tgc Cys 530	cgc Arg	cga Arg	gcc Ala	tac Tyr	aaa Lys 535	att Ile	ctc Leu	acc Thr	gaa Glu	gag Glu 540	gtg Val	ggt Gly	ttc Phe	ccg Pro	1632
ccg Pro 545	gaa Glu	gac Asp	atc Ile	atc Ile	ttc Phe 550	gac Asp	ccg Pro	aat Asn	atc Ile	ttc Phe 555	gcc Ala	gtg Val	gcg Ala	acc Thr	ggt Gly 560	1680
att Ile	gaa Glu	gag Glu	cac His	aac Asn 565	aac Asn	tac Tyr	gcg Ala	cag Gln	gac Asp 570	ttt Phe	atc Ile	ggc Gly	gct Ala	tgc Cys 575	gaa Glu	1728
gac Asp	atc Ile	aaa Lys	cgc Arg 580	gag Glu	ctg Leu	ccg Pro	cac His	gcg Ala 585	ctg Leu	atc Ile	tcc Ser	ggc	ggc Gly 590	gtg Val	tct Ser	1776
					ttc Phe											1824
cac His	gcg Ala 610	gta Val	ttc Phe	ctc Leu	tac Tyr	tac Tyr 615	gcc Ala	atc Ile	cgc Arg	aac Asn	ggt Gly 620	atg Met	gac Asp	atg Met	ggc Gly	1872
					cag Gln 630											1920
ctg Leu	cgc Arg	gat Asp	gcg Ala	gtt Val 645	gaa Glu	gat Asp	gtc Val	att Ile	ctt Leu 650	aac Asn	cgt Arg	cgc Arg	gat Asp	gac Asp 655	ggc Gly	1968
act Thr	gag Glu	cgt Arg	ttg Leu 660	ctg Leu	gat Asp	ttg Leu	gcg Ala	gag Glu 665	aaa Lys	tac Tyr	cgc Arg	ggc Gly	agc Ser 670	aaa Lys	acc Thr	2016
gac Asp	gaa Glu	gct Ala 675	gcc Ala	aac Asn	gcc Ala	cag Gln	cag Gln 680	gcg Ala	gaa Glu	tgg Trp	cgt Arg	agc Ser 685	tgg Trp	gac Asp	gtg Val	2064
aaa Lys	aag Lys	cgt Arg	ctc Leu	gaa Glu	tac Tyr	tcg Ser	ctg Leu	gtg Val	aaa Lys	ggc Gly	att Ile	acc Thr	gaa Glu	ttt Phe	atc Ile	2112

690 695 700

gaa Glu 705	Gln	gat Asp	acc Thr	gaa Glu	gaa Glu 710	gcc Ala	cgt Arg	cag Gln	cag Gln	gcc Ala 715	gcc Ala	cgc Arg	ccg Pro	att Ile	gag Glu 720	2160
gtg Val	att Ile	gaa Glu	999 Gly	ccg Pro 725	ttg Leu	atg Met	gac Asp	ggc Gly	atg Met 730	aac Asn	gtg Val	gtc Val	ggc Gly	gac Asp 735	ctg Leu	2208
ttc Phe	ggc Gly	gaa Glu	999 Gly 740	aaa Lys	atg Met	ttc Phe	ctg Leu	ccg Pro 745	cag Gln	gtg Val	gtg Val	aaa Lys	tcc Ser 750	gct Ala	cgc Arg	2256
gtg Val	atg Met	aaa Lys 755	caa Gln	gcg Ala	gtg Val	gcc Ala	tac Tyr 760	ctg Leu	gag Glu	ccg Pro	ttt Phe	att Ile 765	gaa Glu	gcc Ala	agc Ser	2304
aaa Lys	gaa Glu 770	aaa Lys	ggc Gly	tcc Ser	agc Ser	aac Asn 775	ggc Gly	aag Lys	atg Met	gtg Val	atc Ile 780	gcc Ala	acc Thr	gtg Val	aag Lys	2352
Gly 785	qaA	Val	His	gat Asp	Ile 790	Gly	Lys	Asn	Ile	Val 795	gly	Val	Val	Leu	Gln 800	2400
tgt Cys	aac Asn	aac Asn	tac Tyr	gaa Glu 805	atc Ile	gtc Val	gat Asp	ctt Leu	ggc Gly 810	gtg Val	atg Met	gtg Val	cca Pro	gcg Ala 815	gag Glu	2448
Lys	Ile	Leu	Arg 820	acg Thr	Ala	Arg	Glu	Val 825	Asn	Ala	Asp	Leu	Ile 830	Gly	Leu	2496
Ser	Gly	Leu 835	Ile	acc Thr	Pro	Ser	Leu 840	Asp	Glu	Met	Val	Asn 845	Val	Ala	Lys	2544
Glu	Met 850	Glu	Arg	cag Gln	Gly	Phe 855	Thr	Ile	Pro	Leu	Leu 860	Ile	Gly	Gly	Ala	25 92
Thr 865	Thr	Ser	Lys	gcg Ala	His 870	Thr	Ala	Val	ГÀв	Ile 875	Glu	Gln	Asn	Tyr	Ser 880	2640
Gly	Pro	Thr	Val	tac Tyr 885	Val	Gln	Asn	Ala	Ser 890	Arg	Thr	Val	Gly	Val 895	Val	2688
gcg Ala	gcg Ala	cta Leu	ctc Leu 900	tcc Ser	gac Asp	acc Thr	cag Gln	cgt Arg 905	gat Asp	gac Asp	ttt Phe	gtc Val	gcc Ala 910	cgt Arg	acc Thr	2736
cgc Arg	aaa Lys	gag Glu 915	tac Tyr	gaa Glu	acc Thr	gtg Val	cgt Arg 920	att Ile	cag Gln	cac His	gcc Ala	cgc Arg 925	aaa Lys	aaa Lys	ccg Pro	2784
cgc Arg	acg Thr 930	ccg Pro	ccg Pro	gtc Val	acg Thr	ctg Leu 935	gag Glu	gcg Ala	gcg Ala	cgc Arg	gat Asp 940	aac Asn	gat Asp	ctg Leu	gca Ala	2832

		92		
	c tac acc ccg ccg g Tyr Thr Pro Pro 950			2880
cag gag gtg gaa go Gln Glu Val Glu Al 96				2928
acg ccg ttc ttt at Thr Pro Phe Phe Me 980	g acc tgg tcg ctg t Thr Trp Ser Leu 985		• •	2976
ctg gaa gat gag gt Leu Glu Asp Glu Va 995				3024
gcc aat gat atg ct Ala Asn Asp Met Le 1010				3072
cgt ggc gtg gtg gg Arg Gly Val Val Gl 1025				3120
gaa atc tat cgc ga Glu Ile Tyr Arg As 104	•			3168
cac ctg cgc cag ca His Leu Arg Gln Gl 1060	-	Gly Phe Ala Asn		3216
gcg gat ttt gtc gc Ala Asp Phe Val Al 1075				3264
gct ttc gcg gtg ac Ala Phe Ala Val Th 1090				3312
1105	s Asp Asp Tyr Asn 1110	Lys Ile Met Val 1115	Lys Ala Ile 1120	3360
gcc gac cgt ctg gc Ala Asp Arg Leu Al 112	a Glu Ala Phe Ala			3408
cgt aag gtt tac tg Arg Lys Val Tyr Tr 1140	g gga tat gcg ccg p Gly Tyr Ala Pro 1145	Asn Glu Ser Leu	J . J	3456
1155	u Asn Tyr Gln Gly 1160	Ile Arg Pro Ala 1165	Pro Gly Tyr	3504
cct gcc tgc ccg ga Pro Ala Cys Pro Gl 1170	a cat acc gaa aaa u His Thr Glu Lys 1175	ggc act atc tgg Gly Thr Ile Trp 1180	cag cta ctg Gln Leu Leu	3552
gat gtc gaa aaa ca Asp Val Glu Lys Hi				3600

WO 03/087386 PCT/EP03/04010 93

1190

1195

1200

tgg cca ggc gcg tcg gtc tcc ggc tgg tac ttc agc cat cct gag agc Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His Pro Glu Ser 1205 1210

aaa tac ttc gcg gta gcg cag atc caa cgc gat cag gtg aca gat tat 3696 Lys Tyr Phe Ala Val Ala Gln Ile Gln Arg Asp Gln Val Thr Asp Tyr 1220 1225 1230

gct ttc cgt aaa gga atg agc gtt gag gat gtt gag cgg tgg ctc gcg 3744 Ala Phe Arg Lys Gly Met Ser Val Glu Asp Val Glu Arg Trp Leu Ala 1240

ccg aac ctg ggt tac gat gcg gac tga Pro Asn Leu Gly Tyr Asp Ala Asp 1250 1255

3771

<210> 22

1185

<211> 1256

<212> PRT

<213> Salmonella typhimurium

<400> 22

Met Ser His Val Ala Arg Cys Ser Leu Phe Arg Gln His Ala Leu Cys

Gln Tyr Gly Ser Leu Arg Gly Ala Leu Ser Gly Ala Ser Val Ser Ser

Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile Leu Val Leu

Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu His Glu Glu

Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp Leu Lys Gly

Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile Ala Ala Ile

His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr 100 105

Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Arg Met Glu Ser Leu 120

Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala

Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Phe Val Ala Gly

Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp Val Asn

Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gln Leu Val Ala Ala Tyr 185 190

Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp Leu Ile Leu

195 200 205 Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Val Phe Ala 215 220 Val Lys Glu Glu Phe Glu Ala Leu Gly Val Asp Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr 250 Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala Leu Thr Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln Tyr Val Gln 280 285 Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala Asp Thr Met 315 Ala Lys Gln Ile Arg Glu Trp Ala Glu Ala Gly Phe Leu Asn Ile Val 330 Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala Ala Met Ser Arg Ala Val Ala Gly Leu Leu Pro Arg Gln Leu Pro Asp Ile Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Asp Asp Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Lys Phe 385 390 Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu Asp Val Ala 410 Arg Gln Gln Val Glu Ser Gly Ala Gln Ile Ile Asp Ile Asn Met Asp 425 Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe Leu Ser Leu Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met Ile Asp Ser Ser Lys Trp Glu Val Ile Glu Lys Gly Leu Lys Cys Ile Gln Gly Lys 470 Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Glu Ala Phe Ile 490 His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Met

Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Ala Arg Lys Ile Glu 515 520 525

$\boldsymbol{\alpha}$	_
•	~
·	v

- Ile Cys Arg Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val Gly Phe Pro 530 535 540
- Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala Thr Gly 555 550 560
- Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly Ala Cys Glu 565 570 575
- Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly Gly Val Ser 580 585 590
- Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg Glu Ala Ile 595 600 605
- His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met Asp Met Gly
 610 620
- Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asp Leu Pro Ala Glu 625 630 635 640
- Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg Asp Asp Gly
 645 650 655
- Thr Glu Arg Leu Asp Leu Ala Glu Lys Tyr Arg Gly Ser Lys Thr
 660 665 670
- Asp Glu Ala Ala Asn Ala Gln Gln Ala Glu Trp Arg Ser Trp Asp Val 675 680 685
- Lys Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr Glu Phe Ile 690 695 700
- Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Ile Glu 705 710 715 720
- Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu
 725 730 735
- Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg
 740 745 750
- Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile Glu Ala Ser 755 760 765
- Lys Glu Lys Gly Ser Ser Asn Gly Lys Met Val Ile Ala Thr Val Lys 770 780
- Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln 785 790 795 800
- Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val Pro Ala Glu 805 810 815
- Lys Ile Leu Arg Thr Ala Arg Glu Val Asn Ala Asp Leu Ile Gly Leu 820 825 830
- Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Lys 835 840 845
- Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile Gly Gly Ala 850 855 860

- Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln Asn Tyr Ser 865 870 870 880
- Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val Gly Val Val 885 890 895
- Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val Ala Arg Thr 900 905 910
- Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Ala Arg Lys Lys Pro 915 920 925
- Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn Asp Leu Ala 930 935 940
- Phe Asp Trp Glu Arg Tyr Thr Pro Pro Val Ala His Arg Leu Gly Val 945 950 955 960
- Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr Ile Asp Trp 965 970 975
- Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr Pro Arg Ile 980 985 990
- Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu Phe Lys Asp 995 1000 1005
- Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Leu Leu Asn Pro 1010 1015 1020
- Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly Asp Asp Ile 1025 1030 1035 1040
- Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Leu Thr Val Ser His 1045 1050 1055
- His Leu Arg Gln Gln Thr Glu Lys Val Gly Phe Ala Asn Tyr Cys Leu 1060 1065 1070
- Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp Tyr Ile Gly 1075 1080 1085
- Ala Phe Ala Val Thr Gly Gly Leu Glu Glu Asp Ala Leu Ala Asp Ala 1090 1095 1100
- Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val Lys Ala Ile 1105 1110 1115 1120
- Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His Glu Arg Val 1125 1130 1135
- Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Ser Leu Ser Asn Asp 1140 1145 1150
- Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr 1155 1160 1165
- Pro Ala Cys Pro Glu His Thr Glu Lys Gly Thr Ile Trp Gln Leu Leu 1170 1180
- Asp Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser Phe Ala Met

	W	O 03/0	8738	6											1	PCT/E
										97						
110					119					1199	-			,	1200	
Tr	Pro	o Gly	y Ala	120	r Vai	l Se	r Gly	y Trj	1210	Phe	e Sei	r His	Pro	Glu 1215		
Lys	з Ту	r Phe	2 Ala 122	a Vai	l Ala	a Glı	n Ile	Glr 1225		J Asp	Glr	ı Va]	Thr 1230		Tyr	
Ala	Phe	2 Arg	Lys	s Gly	y Met	t Sei	Va]		ı Asp	Val	Glu	1245		Lev	Ala	
Pro	1250	ı Lev	ı Gly	/ Ty:	r Ası	1255	_	,								
<21 <21	.0> 2 .1> 3 .2> E	771	onel]	a ty	⁄phi											
<22	1> C 2> (DS 1)		8)												
	0> 2															
atg Met 1	ser	cat His	gtt Val	gcc Ala 5	Arg	tgt Cys	tct Ser	ctt Leu	ttc Phe 10	Arg	cag Gln	cac His	gct Ala	ttg Leu 15	Сув	48
cag Gln	tat Tyr	ggc	tcg Ser 20	Leu	cgt Arg	gga Gly	gcg Ala	ttg Leu 25	Ser	gga Gly	gcg Ala	agt Ser	gtg Val 30	agc Ser	agc Ser	96
aaa Lys	gtt Val	gaa Glu 35	caa Gln	ctg Leu	cgt Arg	gcg Ala	cag Gln 40	tta Leu	aat Asn	gaa Glu	cgt Arg	att Ile 45	ctg Leu	gtg Val	ctg Leu	14
gac Asp	ggc Gly 50	ggt Gly	atg Met	ggc Gly	acc Thr	atg Met 55	atc Ile	cag Gln	agc Ser	tat Tyr	cgt Arg 60	cta Leu	cat His	gaa Glu	gaa Glu	19
gat Asp 65	ttc Phe	cgc Arg	Gly 999	gag Glu	cgc Arg 70	ttt Phe	gcc Ala	gac Asp	tgg Trp	ccc Pro 75	tgc Cys	gac	ctg Leu	aaa Lys	ggc Gly 80	240
aac Asn	aat Asn	gac Asp	ctg Leu	ctg Leu 85	gtc Val	ctc Leu	agc Ser	aag Lys	ccg Pro 90	gag Glu	gtg Val	atc Ile	gcc Ala	gct Ala 95	atc Ile	286
cac His	aac Asn	gcc Ala	tac Tyr 100	ttt Phe	gag Glu	gct Ala	ggc Gly	gcg Ala 105	gat Asp	atc Ile	atc Ile	gaa Glu	acc Thr 110	aac Asn	acc Thr	336
ttt Phe	aac Asn	tcg Ser	aca Thr	acc Thr	att Ile	gcg Ala	atg Met	gcg Ala	gat Asp	tac Tyr	cgg Arg	atg Met	gaa Glu	tcc Ser	ctg Leu	384

tcg gcg gaa att aac tat gcg gcg gcc aaa ctg gcg cgc gcc tgc gcc Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala

	_									98						
ga As; 14:	b eri	a tgg	g ace	g gc	g cga a Arg 150	Thi	e cca	a gaa o Glu	a aaa 1 Lys	Pro	Arg	ttt g Phe	gti Val	gcg l Ala	ggc Gly 160	. 480
gte Va	g cti l Lei	ggt Gly	cca Pro	a act	r Asr	cgo Arg	ace Thi	g gcc : Ala	tcc Ser 170	Ile	tcg Ser	p ccg	y gad	gto Val	aac Asn	528
Asp	e ccg	g gcg	ttt Phe 180	Arg	t aat g Asn	ato Ile	acc Thr	Phe 185	: Asp	cag Gln	ctg Leu	gtg Val	gcg Ala 190	Ala	tac Tyr	576
cgt Arg	gaa Glu	Ser 195	The	aaa Lys	gcg Ala	ctg Leu	gtg Val 200	Glu	ggc Gly	Gly 999	gcg Ala	gac Asp 205	Lev	ato Ile	ctg Leu	624
att Ile	gaa Glu 210	Int	gtc Val	Phe	gac Asp	acc Thr 215	Leu	aac Asn	gcc	aaa Lys	gcg Ala 220	Ala	gtg Val	ttt Phe	gcg Ala	672
gtg Val 225	гλя	gaa Glu	gag Glu	Phe	gaa Glu 230	Ala	ctg Leu	ggc	gtt Val	gat Asp 235	Leu	ccg Pro	atc Ile	atg Met	att Ile 240	720
ser	gıy	Thr	116	7hr 245	' Asp	Ala	Ser	Gly	Arg 250	Thr	Leu	Ser	Gly	Gln 255		768
THE	GIU	Ala	260	Tyr	aac Asn	Ser	Leu	Arg 265	His	Ala	Glu	Ala	Leu 270	Thr	Phe	816
GIY	ren	275	Cys	ATA	ctg Leu	GIÀ	Pro 280	Asp	Glu	Leu	Arg	Gln 285	Tyr	Val	Gln	864
gaa Glu	ctg Leu 290	tcg Ser	cgg Arg	att Ile	gcc Ala	gaa Glu 295	tgc Cys	tac Tyr	gtc Val	acc Thr	gcg Ala 300	cac His	ccg Pro	aac Asn	gcc Ala	912
ggc Gly 305	ctg Leu	ccg Pro	aac Asn	gct Ala	ttc Phe 310	ggc Gly	gag Glu	tac Tyr	gac Asp	ctc Leu 315	gac Asp	gcc Ala	gac Asp	acc Thr	atg Met 320	960
VIG	пÅв	GIN	116	325	gaa Glu	Trp	Ala	Glu	Ala 330	Gly	Phe	Leu	Asn	11e 335	Val	1008
GIY	GIY	Сув	340	GIÀ	acc Thr	TNY	Pro	345	His	Ile	Ala	Ala	Met 350	Ser	Arg	1056
gcc Ala	val	gcc Ala 355	ggt Gly	ttg Leu	tcg Ser	Pro	cgc Arg 360	cag Gln	ctg Leu	ccg Pro	gat Asp	atc Ile 365	ccg Pro	gtg Val	gcc Ala	1104
tgc Cys	cgc Arg 370	ctt Leu	tcc Ser	ggc	ctg Leu	gag Glu 375	ccg Pro	ctg Leu	aac Asn	Ile	ggt Gly 380	gac Asp	gat Asp	agc Ser	ctg Leu	1152
ttt Phe	gtc Val	aac Asn	gt <i>c</i> Val	ggc Gly	gaa Glu	cgt Arg	act Thr	aac Asn	gtc Val	acc Thr	ggc Gly	tcg Ser	gcc Ala	aaa Lys	ttt Phe	1200

385	,				390)				99 395	;				400	
aaa Lys	cgc	tto Lev	g ato	aaa Lys 405	Glu	gag Glu	aaa Lys	tac Tyr	agc Ser 410	Glu	gcg Ala	ctg Leu	gat Asp	gtc Val 415	gcc Ala	1248
cgt Arg	cag Gln	Cag	gtc Val 420	Glu	ago Ser	ggc	gcg	Gln 425	Ile	att Ile	gat Asp	atc Ile	aat Asn 430	Met	gat Asp	1296
gag Glu	Gly 999	Atg Met 435	Leu	gac Asp	gcc Ala	gaa Glu	gcg Ala 440	Ala	atg Met	gtg Val	cgt	ttc Phe 445	ctc Leu	agc Ser	ctg Leu	1344
att Ile	gcc Ala 450	Gly	gag Glu	ccg Pro	gac Asp	att Ile 455	gcc Ala	cgt Arg	gta Val	cca Pro	atc Ile 460	Met	att Ile	gac Asp	tcc Ser	1392
tcc Ser 465	aaa Lys	tgg Trp	gag Glu	gtt Val	atc Ile 470	gaa Glu	aaa Lys	61 y 99 g	ctg Leu	aag Lys 475	tgc Cys	att Ile	cag Gln	ggt	aaa Lys 480	1440
ggc	atc Ile	gtc Val	aac Asn	tct Ser 485	att Ile	tcg Ser	atg Met	aaa Lys	gag Glu 490	ggc Gly	gtg Val	gaa Glu	gcc Ala	ttt Phe 495	att Ile	1488
cat His	cat His	gcg Ala	aag Lys 500	ttg Leu	cta Leu	cgt Arg	cgc Arg	tac Tyr 505	ggc Gly	gcc Ala	gca Ala	gtg Val	gtg Val 510	gtg Val	atg Met	1536
gct Ala	ttt Phe	gat Asp 515	gag Glu	cag Gln	999 999	cag Gln	gcc Ala 520	gac Asp	acc Thr	cgc Arg	gaa Glu	cgt Arg 525	aaa Lys	atc Ile	gag Glu	1584
att Ile	tgc Сув 530	cgc Arg	cgc Arg	gct Ala	tac Tyr	aaa Lys 535	att Ile	ttg Leu	ctc Leu	gaa Glu	gag Glu 540	gta Val	ggc Gly	ttt Phe	ccg Pro	1632
ccg Pro 545	gaa Glu	gac Asp	atc Ile	atc Ile	ttc Phe 550	gac Asp	ccg Pro	aat Asn	atc Ile	ttc Phe 555	gcc Ala	gtc Val	gcc Ala	acc Thr	ggt Gly 560	1680
att Ile	gaa Glu	gag Glu	cac His	aac Asn 565	aac Asn	tac Tyr	gcg Ala	cag Gln	gac Asp 570	ttt Phe	atc Ile	ggc Gly	gct Ala	tgt Cys 575	gaa Glu	1728
gac Asp	atc Ile	aaa Lys	cgc Arg 580	gag Glu	ctg Leu	ccg Pro	cac His	gcg Ala 585	ctg Leu	atc Ile	tcc Ser	ggc Gly	ggc Gly 590	gtg Val	tct Ser	1776
aac Asn	gtg Val	tcc Ser 595	ttc Phe	tcg Ser	ttt Phe	cgc Arg	ggc Gly 600	aac Asn	gac Asp	ccg Pro	gta Val	cgt Arg 605	gag Glu	gct Ala	atc Ile	1824
His	gcg Ala 610	gta Val	ttc Phe	ctc Leu	Tyr	tac Tyr 615	gcc Ala	atc Ile	cgc Arg	aac Asn	ggc Gly 620	atg Met	gac Asp	atg Met	ggc Gly	1872
atc Ile 625	gtc Val	aac Asn	gcc Ala	999 Gly	caa Gln 630	ctg Leu	gcg Ala	att Ile	Tyr	gac Asp 635	aac Asn	ctg Leu	cct Pro	gcc Ala	gaa Glu 640	1920

									1	00						
ctg Leu	cgc Arg	gat Asp	gca Ala	yal Val		gat Asp	gtc Val	att Ile	ctt Leu 650	Asn	cgt Arg	cgc Arg	gat Asp	gac Asp 655	ggc Gly	1968
acc Thr	gag Glu	cgt	ttg Leu 660	Leu	gat Asp	ttg Leu	gcg Ala	gag Glu 665	aaa Lys	tat Tyr	cgc Arg	ggc	agc Ser 670	Lys	acc Thr	2016
gac Asp	gag Glu	gct Ala 675	Ala	agt Ser	gcc	cag Gln	cag Gln 680	Ala	gaa Glu	tgg Trp	cgt Arg	agc Ser 685	tgg Trp	gac Asp	gtg Val	2064
aaa Lys	aag Lys 690	cgt Arg	ctc Leu	gaa Glu	tac Tyr	tcg Ser 695	ctg Leu	gtg Val	aaa Lys	ggc	att Ile 700	acc Thr	gag Glu	ttt Phe	atc Ile	2112
gaa Glu 705	cag Gln	gat Asp	acc Thr	gaa Glu	gaa Glu 710	gcc Ala	cgt Arg	cag Gln	cag Gln	gcc Ala 715	gcc Ala	cgc Arg	ccg Pro	att Ile	gag Glu 720	2160
gtg Val	att Ile	gaa Glu	61 y 999	ccg Pro 725	ctg Leu	atg Met	gac Asp	ggc Gly	atg Met 730	aac Asn	gtg Val	gtc Val	ggc Gly	gac Asp 735	ctg Leu	2208
ttc Phe	ggc	gaa Glu	999 Gly 740	aaa Lys	atg Met	ttc Phe	ctg Leu	ccg Pro 745	cag Gln	gtg Val	gtg Val	aaa Lys	tcc Ser 750	gct Ala	cgc Arg	2256
gtg Val	atg Met	aaa Lys 755	caa Gln	gcg Ala	gtg Val	gcc Ala	tac Tyr 760	ctg Leu	gag Glu	ccg Pro	ttt Phe	att Ile 765	gaa Glu	gcc	agc Ser	2304
Lys	Glu 770	Lys	Gly	Ser	agc Ser	Asn 775	Gly	Lys	Met	Val	Ile 780	Ala	Thr	Val	Lys	2352
ggc Gly 785	gat Asp	gtg Val	cac His	gac Asp	att Ile 790	ggc	aag Lys	aac Asn	att Ile	gtc Val 795	ggc Gly	gtg Val	gtg Val	ctg Leu	caa Gln 800	2400
Сув	Asn	Asn	Tyr	Glu 805	atc Ile	Val	qaA	Leu	Gly B10	Val	Met	Val	Pro	Ala 815	Glu	2448
aaa Lys	atc Ile	ctc Leu	aga Arg 820	acg Thr	gcg Ala	cgt Arg	gaa Glu	gtg Val 825	aat Asn	gcc Ala	gat Asp	ctg Leu	att Ile 830	ggt Gly	ctt Leu	2496
tcc Ser	ggg Gly	ctt Leu 835	atc Ile	acc Thr	ccg Pro	tcg Ser	ctg Leu 840	gac Asp	gaa Glu	atg Met	gtc Val	aac Asn 845	gtg Val	gcg Ala	aaa Lys	2544
gag Glu	atg Met 850	gag Glu	cgt Arg	cag Gln	ggc Gly	ttt Phe 855	act Thr	atc Ile	ccg Pro	cta Leu	ctg Leu 860	atc Ile	ggc Gly	ggc Gly	gca Ala	2592
acc Thr 865	act Thr	tcg Ser	aaa Lys	gcg Ala	cat His 870	acg Thr	gcg Ala	gtg Val	aaa Lys	atc Ile 875	gag Glu	cag Gln	aac Asn	tac Tyr	agc Ser 880	2640
ggt Gly	ccg Pro	acg Thr	gtc Val	tac Tyr	gtg Val	cag Gln	aat Asn	gct Ala	tcg Ser	cgt Arg	acc Thr	gtg Val	ggc Gly	gtg Val	gtg Val	2688

1	885	890	895
gcg gcg cta ctc (Ala Ala Leu Leu (900	tee gae ace cag egt Ser Asp Thr Gln Arg 905	gat gac ttt gtc gcc g Asp Asp Phe Val Ala 910	Arg Thr
cgc aaa gag tac g Arg Lys Glu Tyr (915	gaa acc gtg cgt att Glu Thr Val Arg Ile 920	cag cac gcc cgc aaa Gln His Ala Arg Lys 925	aaa ccg 2784 Lys Pro
cgc acg ccg ccg g Arg Thr Pro Pro V 930	gtc acg ctg gaa gcg /al Thr Leu Glu Ala 935	g gcg cgc gat aat gat Ala Arg Asp Asn Asp 940	ctg gca 2832 Leu Ala
ttt gat tgg gaa o Phe Asp Trp Glu A 945	gc tac acc ccg ccg urg Tyr Thr Pro Pro 950	gta gcc cac cgt ctg Val Ala His Arg Leu 955	ggc gtg 2880 Gly Val 960
Gln Glu Val Glu A	da Ser Ile Glu Thr 165	ctg cgc aac tac atc Leu Arg Asn Tyr Ile 970	Asp Trp 975
Thr Pro Phe Phe M 980	et Thr Trp Ser Leu 985		Arg Ile
Leu Glu Asp Glu V 995	al Val Gly Val Glu 1000	gcg cag cgt ctg ttt Ala Gln Arg Leu Phe 1005	Lys Asp
Ala Asn Asp Met L 1010	eu Asp Lys Leu Ser 1015	gcc gag aaa ctg ttg Ala Glu Lys Leu Leu 1020	Asn Pro
Arg Gly Val Val G 1025	ly Leu Phe Pro Ala 1030	aac cgt gtg ggt gac Asn Arg Val Gly Asp 1035	Asp Ile 1040
gaa atc tat cgc g Glu Ile Tyr Arg A 10	sp Glu Thr Arg Thr	cat gtt ctg acg gtc His Val Leu Thr Val 1050	agc cac 3168 Ser His 1055
cac ctg cgc cag c His Leu Arg Gln G 1060	ag acc gag aaa gtt ln Thr Glu Lys Val 1065	ggt ttt gct aac tac Gly Phe Ala Asn Tyr 1070	Cys Leu
Ala Asp Phe Val A	la Pro Lys Leu Ser 1080	ggc aaa gcg gac tac Gly Lys Ala Asp Tyr 1085	Ile Gly
gct ttc gcg gtg ad Ala Phe Ala Val Ti 1090	cc ggc ggt ctg aag nr Gly Gly Leu Lys 1095	gag gat gcg ctg gcg Glu Asp Ala Leu Ala 1100	gac gcc 3312 Asp Ala
ttc gaa gcg caa ca Phe Glu Ala Gln Hi 1105	ac gac gac tat aac is Asp Asp Tyr Asn 1110	aag atc atg gtg aaa Lys Ile Met Val Lys 1115	gcg att 3360 Ala Ile 1120
gcc gac cgt ctg gc Ala Asp Arg Leu Al 112	a Glu Ala Phe Ala	gag tat ctg cat gag Glu Tyr Leu His Glu 130 1	cgt gta 3408 Arg Val 135

102	
cgt aag gtt tac tgg gga tat gcg ccg aac gag agc ctg agt aac gac Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Ser Leu Ser Asn Asp 1140 1145 1150	3456
gaa tta atc cgc gaa aac tac cag ggg att cgc ccg gcg ccg ggt tat Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr 1155 1160 1165	3504
cct gcc tgc ccg gaa cat acc gaa aaa ggc act atc tgg cag cta ctg Pro Ala Cys Pro Glu His Thr Glu Lys Gly Thr Ile Trp Gln Leu Leu 1170 1175 1180	3552
gat gtc gaa aaa cac acc ggg atg aag ctc acc gaa tct ttc gcc atg Asp Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser Phe Ala Met 1185 1190 1195 1200	3600
tgg cct ggc gcg tcg gtc tcc ggc tgg tac ttc agc cat cct gag agc Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His Pro Glu Ser 1205 1210 1215	3648
aaa tac ttc gcg gta gcg cag atc caa cgc gat cag gtg aca gat tat Lys Tyr Phe Ala Val Ala Gln Ile Gln Arg Asp Gln Val Thr Asp Tyr 1220 1225 1230	369 6
gct ttc cgt aaa gga atg agc gtt gag gac gtt gag cgg tgg ctc gcg Ala Phe Arg Lys Gly Met Ser Val Glu Asp Val Glu Arg Trp Leu Ala 1235 1240 1245	3744
ccg aac ctg ggt tac gat gcg gac tga Pro Asn Leu Gly Tyr Asp Ala Asp 1250 1255	3771
<210> 24 <211> 1256 <212> PRT <213> Salmonella typhi	
•	
<400> 24 Met Ser His Val Ala Arg Cys Ser Leu Phe Arg Gln His Ala Leu Cys 1 5 10 15	
Gln Tyr Gly Ser Leu Arg Gly Ala Leu Ser Gly Ala Ser Val Ser Ser 20 25 30	
Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile Leu Val Leu 35 40 45	
Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu His Glu Glu 50 55 60	
Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp Leu Lys Gly 65 70 75 80	
Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile Ala Ala Ile 85 90 95	
His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr 100 105 110	

- Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala 130 135 140
- Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Phe Val Ala Gly
 145 150 155 160
- Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp Val Asn 165 170 175
- Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gln Leu Val Ala Ala Tyr 180 185 190
- Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp Leu Ile Leu 195 200 205
- Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Val Phe Ala 210 215 220
- Val Lys Glu Glu Phe Glu Ala Leu Gly Val Asp Leu Pro Ile Met Ile 225 230 235 240
- Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr 245 250 255
- Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala Leu Thr Phe 260 265 270
- Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln Tyr Val Gln 275 280 285
- Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His Pro Asn Ala 290 295 300
- Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala Asp Thr Met 305 310 315 320
- Ala Lys Gln Ile Arg Glu Trp Ala Glu Ala Gly Phe Leu Asn Ile Val
- Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala Ala Met Ser Arg 340 345 350
- Ala Val Ala Gly Leu Ser Pro Arg Gln Leu Pro Asp Ile Pro Val Ala 355 360 365
- Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Asp Asp Ser Leu 370 380
- Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Lys Phe 385 390 395 400
- Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu Asp Val Ala 405 410 415
- Arg Gln Gln Val Glu Ser Gly Ala Gln Ile Ile Asp Ile Asn Met Asp
 420 425 430
- Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe Leu Ser Leu 435 440 445
- Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met Ile Asp Ser

450

455

460

Ser Lys Trp Glu Val Ile Glu Lys Gly Leu Lys Cys Ile Gln Gly Lys 465 470 475 480

Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Glu Ala Phe Ile 485 490 495

His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Met
500 505 510

Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Glu Arg Lys Ile Glu 515 520 525

Ile Cys Arg Arg Ala Tyr Lys Ile Leu Leu Glu Glu Val Gly Phe Pro 530 535 540

Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala Thr Gly 545 550 555 555

Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly Ala Cys Glu 565 570 575

Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly Gly Val Ser 580 585 590

Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg Glu Ala Ile 595 600 605

His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met Asp Met Gly 610 620

Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asn Leu Pro Ala Glu 625 635 640

Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg Asp Asp Gly 645 655

Thr Glu Arg Leu Leu Asp Leu Ala Glu Lys Tyr Arg Gly Ser Lys Thr
660 665 670

Asp Glu Ala Ala Ser Ala Gln Gln Ala Glu Trp Arg Ser Trp Asp Val 675 680 685

Lys Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr Glu Phe Ile 690 695 700

Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Ile Glu 705 710 715 720

Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu
725 730 735

Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg
740 745 750

Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile Glu Ala Ser 755 760 765

Lys Glu Lys Gly Ser Ser Asn Gly Lys Met Val Ile Ala Thr Val Lys 770 785

WO 03/087386 PCT/EP03/04010

4		
7	115	
- 1	uii	

- Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln
 785 790 795 800
- Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val Pro Ala Glu 805 810 815
- Lys Ile Leu Arg Thr Ala Arg Glu Val Asn Ala Asp Leu Ile Gly Leu 820 825 830
- Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Lys
- Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile Gly Gly Ala 850 855 860
- Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln Asn Tyr Ser 865 870 875 880
- Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val Gly Val Val 885 890 895
- Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val Ala Arg Thr 900 905 910
- Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Ala Arg Lys Lys Pro 915 920 925
- Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn Asp Leu Ala 930 935 940
- Phe Asp Trp Glu Arg Tyr Thr Pro Pro Val Ala His Arg Leu Gly Val 945 950 955 960
- Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr Ile Asp Trp 965 970 975
- Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr Pro Arg Ile 980 985 985
- Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu Phe Lys Asp 995 1000 1005
- Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Leu Leu Asn Pro 1010 1015 1020
- Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly Asp Asp Ile 1025 1030 1035 1040
- Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Leu Thr Val Ser His
 1045 1050 1055
- His Leu Arg Gln Gln Thr Glu Lys Val Gly Phe Ala Asn Tyr Cys Leu 1060 1065 1070
- Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp Tyr Ile Gly 1075 1080 1085
- Ala Phe Ala Val Thr Gly Gly Leu Lys Glu Asp Ala Leu Ala Asp Ala 1090 1095 1100
- Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val Lys Ala Ile 1105 1110 1115 1120

									•	106				•		
Ala	a Ası) Arg	g Lev	1 Ala 1125	Glu G	ı Ala	Phe	e Ala	Glu 1130		Lev	His	Glu	Arg 1135	Val	•
Arg	j Lys	Val	Туз 1140	r Trp	Gly	/ Туг	Ala	Pro 1145		Glu	Ser	Leu	Ser 1150		Asp	
Glu	ı Leı	lle 1155	Arg	g Glu	Asr	туг	Glr 1160		' Ile	Arg	Pro	Ala 1165		Gly	Tyr	
Pro	Ala 1170	Cys	Pro	Glu	His	Thr 1175		Lys	Gly		Ile 1180	-	Gln	Leu	Leu	
Asp 118	Val	Glu	Lys	His	Thr 1190		Met	Lye		Thr 1195		Ser	Phe		Met 1200	
Trp	Pro	Gly	Ala	Ser 1205	Val	Ser	Gly		Tyr 1210		Ser	His		Glu 1215	Ser	
Lys	Tyr	Phe	Ala 1220	Val	Ala	Gln	Ile	Gln 1225		Asp	Gln		Thr 1230		Tyr	
Ala	Phe	Arg 1235	Lys	Gly	Met		Val 1240		Asp	Val		Arg 1245		Leu	Ala	
	Asn 1250		Gly	Tyr		Ala 1255	Asp	ı					•	٠,		
<21	0> 2	5											'			
<21	1> 3	711												,	•	
	2> D 3> P		omon	as f	luor	escei	ns							,		
<22	1> C 2> (1)	-	8)										•		
< 2 2 .	3> K.	PU03!	563						•							
atg	0> 2: tcc Ser	gat	cgc	agc Ser 5	gtc Val	cgc Arg	ctt Leu	caa Gln	gct Ala 10	ctc Leu	aag Lys	caa Gln	gct Ala	ctc Leu 15	aaa Lys	48
gag Glu	cgc Arg	atc Ile	ctg Leu 20	att Ile	ctc Leu	gac Asp	ggc Gly	ggc Gly 25	atg Met	ggc Gly	acg Thr	atg Met	atc Ile 30	cag Gln	agc Ser	96
tac Tyr	aag Lys	ctc Leu 35	gaa Glu	gag Glu	cag Gln	gat Asp	tat Tyr 40	cgc Arg	ggc Gly	aaa Lys	cgc Arg	ttc Phe 45	gcc Ala	gac Asp	tgg Trp	144
ccg Pro	agc Ser 50	gac Asp	gtc Val	aag Lys	ggc Gly	aac Asn 55	aac Asn	gac Asp	ctg Leu	ttg Leu	gtg Val 60	ctg Leu	acc Thr	cgc Arg	ccg Pro	192
gac Asp 65	gtg Val	atc Ile	ggc Gly	ggc Gly	atc Ile 70	gag Glu	aaa Lys	gcc Ala	tat Tyr	ctg Leu 75	gat Asp	gcc Ala	ggt Gly	gcc Ala	gac Asp 80	240

atc ctc gag acc aac acc ttc aac gcc acg cag att tcc atg gcc gac Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Ile Ser Met Ala Asp

107 90

				63	1				90					35		
tac Tyr	ggc	atg Met	gaa Glu 100	Glu	ctg Leu	gtc Val	tac Tyr	gaa Glu 105	ctc Leu	aac Asn	gta Val	gaa Glu	ggc Gly 110	gcc Ala	cgt Arg	336
ctg Leu	gca Ala	cgc Arg 115	aag Lys	gtc Val	gcc Ala	gac Asp	gcg Ala 120	aaa Lys	acc Thr	ctc Leu	gag Glu	acc Thr 125	ccc Pro	gac Asp	aag Lys	384
ccg Pro	cgc Arg 130	ttc Phe	gtc Val	gcc Ala	ggc	gtt Val 135	ctc Leu	ggc	ccg Pro	acc Thr	agc Ser 140	cgc Arg	acc Thr	tgc Cys	tcg Ser	432
ctg Leu 145	tcg Ser	ccg Pro	gac Asp	gtc Val	aac Asn 150	aac Asn	ccg Pro	ggc Gly	tat Tyr	cgc Arg 155	aac Asn	gtc Val	acc Thr	ttc Phe	gat Asp 160	480
gag Glu	ctg Leu	gtc Val	gaa Glu	aac Asn 165	tac Tyr	acc Thr	gag Glu	gcc Ala	acc Thr 170	aaa Lys	gge Gly	ctg Leu	atc Ile	gag Glu 175	ggc	528
ggc Gly	gcg Ala	gat Asp	ctg Leu 180	atc Ile	ctg Leu	atc Ile	gaa Glu	acc Thr 185	atc Ile	ttc Phe	gac Asp	acc Thr	ctc Leu 190	aac Asn	gcc Ala	576
aaa Lys	gcc Ala	gcg Ala 195	atc Ile	ttc Phe	gcc Ala	gtg Val	caa Gln 200	ggc Gly	gtg Val	ttc Phe	gaa Glu	gaa Glu 205	ctg Leu	ggc Gly	ttc Phe	624
gaa Glu	ttg Leu 210	ccg Pro	atc Ile	atg Met	atc Ile	tcc Ser 215	ggc Gly	acc Thr	atc Ile	acc Thr	gac Asp 220	Ala	tcc Ser	ggc Gly	cgt Arg	672
Thr 225	Leu	Ser	Gly	Gln	acc Thr 230	Thr	Glu	Ala	Phe	Trp 235	Asn	Ser	Val	Ala	His 240	720
gcc Ala	aaa Lys	ccg Pro	att Ile	tcc Ser 245	gtc Val	ggt Gly	ctt Leu	aac Asn	tgc Cys 250	gcc Ala	ctc Leu	ggc Gly	gcc Ala	cgc Arg 255	gaa Glu	768
ctg Leu	cgt Arg	ccg Pro	tac Tyr 260	ctg Leu	gaa Glu	gag Glu	ctg Leu	tcg Ser 265	gac Asp	aag Lys	gcc Ala	agc Ser	acc Thr 270	cac His	gtt Val	816
tcg Ser	gcg Ala	cac His 275	ccg Pro	aac Asn	gcc Ala	ggc Gly	ctg Leu 280	ccg Pro	aac Asn	gaa Glu	ttc Phe	ggc Gly 285	gag Glu	tac Tyr	ga c	864
gag Glu	ctg Leu 290	ccg Pro	gtg Val	gac	acc Thr	gcc Ala 295	aag Lys	gtc Val	atc Ile	gaa Glu	gag Glu 300	ttc Phe	gcc Ala	cag Gln	agc Ser	912
ggt Gly 305	ttc Phe	ctc Leu	aac Asn	atc Ile	gtc Val 310	ggc Gly	ggt Gly	tgc Cys	tgc Cys	ggc Gly 315	acc Thr	acg Thr	ccg Pro	Gly	cat His 320	960
atc Ile	gaa Glu	gcc Ala	atc Ile	gcc Ala 325	aaa Lys	gcc Ala	gtt Val	gcc Ala	ggt Gly 330	tac Tyr	gcg Ala	cca Pro	cgg Arg	cag Gln 335	att Ile	1008

									•	108						
Pro	y Asp	att Ile	2 Pro	Lys	gcc Ala	Cys	Arg	ctg Leu 345	Ser	ggt	ctg Leu	gaa Glu	Pro 350	. Phe	acc Thr	1056
att Ile	gat Asp	cgc Arg 355	, Ser	tcg Ser	ctg Leu	ttc Phe	gto Val 360	aac Asn	gtc Val	ggc Gly	gag Glu	cgg Arg 365	acc Thr	aac Asn	atc	1104
acc Thr	999 Gly 370	Ser	gcg Ala	aaa Lys	ttt Phe	gcc Ala 375	cgg	ctg Leu	atc	cgt Arg	gaa Glu 380	gac Asp	aac Asn	tac	acc Thr	1152
gaa Glu 385	Ala	ctg Leu	gaa Glu	gtc Val	gcc Ala 390	ctg Leu	cag Gln	cag Gln	gtc Val	gag Glu 395	Ala	ggc Gly	gcc Ala	cag Gln	gtg Val 400	1200
atc Ile	gac Asp	atc Ile	aac Asn	atg Met 405	gac Asp	gaa Glu	G1y 999	atg Met	ctc Leu 410	gat Asp	tcg Ser	aag Lys	aag Lys	gcc Ala 415	atg Met	1248
gtg Val	acc	ttc Phe	ctc Leu 420	aat Asn	ctg Leu	att Ile	gcc	ggc Gly 425	gaa Glu	ccg Pro	gac Asp	atc Ile	tcc Ser 430	cgc Arg	gta Val	1296
ccg Pro	atc	atg Met 435	atc Ile	gac Asp	tcc Ser	tcg Ser	aaa Lys 440	tgg Trp	gac Asp	gtg Val	atc Ile	gaa Glu 445	gcc Ala	ggc	ctc Leu	1344
aag Lys	tgc Cys 450	att Ile	cag Gln	ggc Gly	aag Lys	ggc Gly 455	atc Ile	gtc Val	aac Asn	tcg Ser	atc Ile 460	agc Ser	atg Met	aaa Lys	gaa Glu	1392
ggc Gly 465	gtc Val	gag Glu	cag Gln	ttc Phe	atc Ile 470	cac His	cac His	gcc Ala	aaa Lys	ctg Leu 475	tgc Сув	aag Lys	cgc Arg	tat Tyr	ggc Gly 480	1440
gcc Ala	gcc Ala	gtg Val	gtg Val	gtg Val 485	atg Met	gcg Ala	ttc Phe) Asp	gaa Glu 490	gcc Ala	ggc	cag Gln	gct Ala	gac Asp 495	acc Thr	1488
gaa Glu	gcg Ala	cgc Arg	aag Lys 500	aaa Lys	gag Glu	atc Ile	tgc Cys	aaa Lys 505	cgc Arg	tcc Ser	tac Tyr	gac Asp	att Ile 510	ctg Leu	gtc Val	1536
aac Asn	gaa Glu	gtc Val 515	ggc Gly	ttc Phe	ccg Pro	ccg Pro	gaa Glu 520	gac Asp	atc Ile	att Ile	ttc Phe	gac Asp 525	ccg Pro	aac Asn	atc Ile	1584
ttc Phe	gcc Ala 530	gtg Val	gcc Ala	acc Thr	ggc Gly	atc Ile 535	gaa Glu	gaa Glu	cac His	aac Asn	aac Asn 540	tac Tyr	gct Ala	gtg Val	gac Asp	1632
ttc Phe 545	atc Ile	aac Asn	gcc Ala	Cys	gcc Ala 550	tac Tyr	atc Ile	cgc Arg	gac Asp	gag Glu 555	ctg Leu	ccg Pro	tat Tyr	gcc Ala	ctg Leu 560	1680
agc Ser	tcc Ser	ggc Gly	Gly	gtg Val 565	tcc Ser	aac Asn	gtg Val	tcg Ser	ttc Phe 570	tcg Ser	ttc Phe	cgc Arg	ggc Gly	aac Asn 575	aac Asn	1728
ccg Pro	gtg Val	cgc Arg	gag Glu	gcg Ala	atc Ile :	cac His	tcg Ser	gtg Val	ttc Phe	ctg Leu	ctg Leu	tac Tyr	gcg Ala	atc Ile	cgc Arg	1776

580 58

585

													550			
gcc Ala	ggc	Leu 595	Thr	atg Met	ggt	atc Ile	gtc Val 600	aac Asn	gcc	ggt	cag Gln	ctg Leu 605	gag Glu	atc Ile	tac Tyr	1824
Asp	Cag Gln 610	Ile	ccg Pro	cag Gln	gaa Glu	ctg Leu 615	Arg	gac Asp	gcc	gtt Val	gaa Glu 620	gac Asp	gtg Val	atc Ile	ctc Leu	1872
aac Asn 625	Arg	acg Thr	ccg Pro	gaa Glu	Gly 630	Thr	gac Asp	gcc	ctc Leu	ctc Leu 635	gcc Ala	atc Ile	gcc Ala	gac Asp	aag Lys 640	1920
Tyr	Lys	Gly	Asp	Gly 645	Ser	Val	Lys	gaa Glu	Ala 650	Glu	Thr	Glu	Glu	Trp 655	Arg	1968
Gly	Trp	Asp	Val 660	Asn	Lys	Arg	Leu	gaa Glu 665	His	Ala	Leu	Val	Lys 670	Gly	Ile	2016
Thr	Thr	His 675	Ile	Val	Glu	Asp	Thr 680	gaa Glu	Glu	Ser	Arg	Gln 685	Ser	Phe	Ala	2064
Arg	Pro 690	Ile	Glu	Val	Ile	Glu 695	Gly	ccg Pro	Leu	Met	Ser 700	Gly	Met	Asn	Ile	2112
Val 705	Gly	Asp	Leu	Phe	Gly 710	Ala	Gly	aaa Lys	Met	Phe 715	Leu	Pro	Gln	Val	Val 720	2160
Lys	Ser	Ala	Arg	Val 725	Met	Lys	Gln	gcc Ala	Val 730	Ala	His	Leu	Ile	Pro 735	Phe	2208
11e	GIU	Leu	740	Lys	Gly	Asp	Lys	ccg Pro 745	Glu	Ala	Lys	Gly	Lys 750	Ile	Leu	2256
Met	Ala	755	Val	Lys	Gly	Asp	Val 760	cac His	Asp	Ile	Gly	Lys 765	Asn	Ile	Val	2304
Gly	Val 770	Val	Leu	Gly	Сув	Asn 775	Gly	tac Tyr	Asp	Ile	Val 780	Asp	Leu	Gly	Val	2352
atg Met 785	gtg Val	ccg Pro	gcg Ala	gag Glu	aag Lys 790	atc Ile	ctg Leu	cag Gln	gtg Val	gcc Ala 795	aag Lys	gag Glu	cag Gln	aag Lys	tgc Cys 800	2400
Asp	atc Ile	atc Ile	ggc	ctg Leu 805	tcc Ser	ggt Gly	ctg Leu	atc Ile	acc Thr 810	ccg Pro	tcg Ser	ctg Leu	gat Asp	gag Glu 815	atg Met	2448
gtc Val	cat His	Val	gcc Ala 820	cgc Arg	gag Glu	atg Met	Gln	cgc Arg 825	cag Gln	gac Asp	ttc Phe	His	ctg Leu 830	ccg Pro	ctg Leu	2496

										110		•		•		
atg Met	at: Il:	e Gl	A GT	c gc	g ac	c acc	Sez 840	Ly	g gcg	g cac a His	ace Thi	gcg Ala 845	Va]	g aag l. Lys	atc Ile	2544
gag Glu	Pro 850	э гу	g tad	c age	c aad r Asi	n Asp 855	Ala	gtg Val	g gto Val	tac Tyr	ytg Val	Thr	gac Aap	gco Ala	tcc Ser	2592
cgc Arg 865	gco	gte a Val	g gg(gte Val	9 909 l Ala 870	a Thr	cag Gln	ttg Lev	, ct <u>c</u> , Lev	Ser 875	. TA8	gaa Glu	ctg	aaa Lys	gcc Ala 880	2640
ggt Gly	Phe	gto Val	c caq l Glr	aag Lys 885	Thi	cgc Arg	gaa Glu	gag	tac Tyr 890	Ile	gac Asp	gtc Val	ċgc Arg	gag Glu 895		2688
acc Thr	gcc	aac Asr	c cgc Arg 900	, ser	gcc Ala	cgc Arg	acc Thr	gaa Glu 905	Arg	ctg Leu	agc Ser	tac Tyr	gcc Ala 910	Ala	gcg	2736
atc Ile	gcc Ala	Lys 915	Lys	Pro	g cag Gln	ttc Phe	gac Asp 920	tgg Trp	gcc Ala	act Thr	tac Tyr	acc Thr 925	ccg	gtc Val	aaa Lys	2784
ccg Pro	acc Thr 930	Pne	acc Thr	Gly	acc Thr	cgc Arg 935	gtg Val	ctg Leu	gac Asp	aac Asn	atc Ile 940) Asp	ctc Leu	aac	gtt Val	2832
ctc Leu 945	gcc Ala	gag Glu	tac Tyr	atc Ile	gac Asp 950	tgg Trp	acg Thr	ccg Pro	ttc Phe	ttc Phe 955	atc Ile	tcc Ser	tgg Trp	gac Asp	ctg Leu 960	2880
gcc Ala	ggc Gly	aag Lys	ttc Phe	ccg Pro 965	Arg	atc Ile	ctc Leu	gaa Glu	gac Asp 970	gaa Glu	gtg Val	gtc Val	ggc	gaa Glu 975	gcg Ala	2928
gcg Ala	acc Thr	gcg Ala	ctg Leu 980	tac Tyr	aag Lys	gac Asp	gct Ala	cgc Arg 985	gag Glu	atg Met	ctg Leu	acc Thr	aag Lys 990	ctg Leu	atc Ile	2976
gac Asp	gag Glu	ааа L ув 995	ctg Leu	atc Ile	agc Ser	gcc Ala 1	cgt Arg .000	gcg Ala	gtg Val	ttc Phe	Gly	ttc Phe 1005	tgg Trp	ccg Pro	gcc Ala	3024
WRII A	cag Gln 010	gtg Val	cac His	gac Asp	qaA	gat Asp 1015	atc Ile	gag Glu	ctg Leu	Tyr	ggc Gly 020	gat Asp	gac Asp	ggc Gly	aag Lys	3072
cca a Pro 1 1025	atg Met	gcg Ala	cgc Arg	ctg Leu	cat His 1030	cac His	ctg Leu	cgc Arg	cag Gln	cag Gln 1035	Ile	atc Ile	aag Lys	acc Thr	gac Asp 1040	3120
ggc a	aaa Lys	ccg Pro	ABD	ttc Phe 045	tcc Ser	ctc Leu	gcc Ala	Asp	ttc Phe 050	gtc Val	gcg Ala	cċg Pro	Lys	gac Asp .055	agc Ser	3168
gaa g Glu N	gtg /al	inr	gac Asp 060	tac Tyr	gtt Val	ggt Gly	Gly	ttc Phe 065	atc Ile	acc Thr	acc Thr	Ala	999 Gly 070	atc Ile	ggc Gly	3216
dec e	gaa Blu	gaa Glu	gtg Val	gcc Ala	aag Lys	gcc Ala '	tat (Tyr (cag Gln	gac Asp	gcc (ggc Gly	gac Asp	gat Asp	tac Tyr	aac Asn	3264

1075

1080

1085

teg ate atg gte aag gee etg gee gae egt etg gee gag geg tge gee Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Cys Ala 1095 1100 gag tgg ctg cac cag cag gtg cgc aaa gag cac tgg ggt tac gcc aag Glu Trp Leu His Gln Gln Val Arg Lys Glu His Trp Gly Tyr Ala Lys 1110 1115 gat gaa gcc ctc gat aac gag gcg ctg atc aaa gag cag tat tcc ggc Asp Glu Ala Leu Asp Asn Glu Ala Leu Ile Lys Glu Gln Tyr Ser Gly 1125 1130 ate ege eet gee eee gge tae eeg geg tge eeg gat eac ace gag aag 3456 Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys 1145 gec ace etg tte gec etg ete gac eet gaa gea eag gaa atg ege gee 3504 Ala Thr Leu Phe Ala Leu Leu Asp Pro Glu Ala Gln Glu Met Arg Ala 1155 1160 ggc cgc agc ggt gtg ttc ctc acc gag cac tac gcg atg ttc ccg gcg 3552 Gly Arg Ser Gly Val Phe Leu Thr Glu His Tyr Ala Met Phe Pro Ala 1170 1175 1180 gca gcc gtc agc ggc tgg tac ttc gcc cat ccg cag gcg cag tac ttc Ala Ala Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe 1185 1190 gcc gtg ggc aag gtc gac aag gat cag gtg cag agc tac acc tcg cgc Ala Val Gly Lys Val Asp Lys Asp Gln Val Gln Ser Tyr Thr Ser Arg 1210 aaa ggc cag gaa ctg agc ctg acc gag cgc tgg ctg gca ccc aat ctg Lys Gly Gln Glu Leu Ser Leu Thr Glu Arg Trp Leu Ala Pro Asn Leu 1220 ggc tac gac aac tga 3711 Gly Tyr Asp Asn 1235 <210> 26 <211> 1236

<212> PRT

<213> Pseudomonas fluorescens

<400> 26

Met Ser Asp Arg Ser Val Arg Leu Gln Ala Leu Lys Gln Ala Leu Lys

Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr Met Ile Gln Ser

Tyr Lys Leu Glu Glu Gln Asp Tyr Arg Gly Lys Arg Phe Ala Asp Trp

Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu Val Leu Thr Arg Pro 50 60

Asp Val Ile Gly Gly Ile Glu Lys Ala Tyr Leu Asp Ala Gly Ala Asp

112 70 Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Ile Ser Met Ala Asp 85 90 95 Tyr Gly Met Glu Glu Leu Val Tyr Glu Leu Asn Val Glu Gly Ala Arg 105 Leu Ala Arg Lys Val Ala Asp Ala Lys Thr Leu Glu Thr Pro Asp Lys 120 1. Pro Arg Phe Val Ala Gly Val Leu Gly Pro Thr Ser Arg Thr Cys Ser 135 Leu Ser Pro Asp Val Asn Asn Pro Gly Tyr Arg Asn Val Thr Phe Asp 155 Glu Leu Val Glu Asn Tyr Thr Glu Ala Thr Lys Gly Leu Ile Glu Gly 170 Gly Ala Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala 180 Lys Ala Ala Ile Phe Ala Val Gln Gly Val Phe Glu Glu Leu Gly Phe Glu Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Thr Glu Ala Phe Trp Asn Ser Val Ala His 225 230 235 Ala Lys Pro Ile Ser Val Gly Leu Asn Cys Ala Leu Gly Ala Arg Glu 250 Leu Arg Pro Tyr Leu Glu Glu Leu Ser Asp Lys Ala Ser Thr His Val 265 Ser Ala His Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp 280 Glu Leu Pro Val Asp Thr Ala Lys Val Ile Glu Glu Phe Ala Gln Ser Gly Phe Leu Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Gly His Ile Glu Ala Ile Ala Lys Ala Val Ala Gly Tyr Ala Pro Arg Gln Ile Pro Asp Ile Pro Lys Ala Cys Arg Leu Ser Gly Leu Glu Pro Phe Thr 345 Ile Asp Arg Ser Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Ile

Thr Gly Ser Ala Lys Phe Ala Arg Leu Ile Arg Glu Asp Asn Tyr Thr

Glu Ala Leu Glu Val Ala Leu Gln Gln Val Glu Ala Gly Ala Gln Val

395

WO 03/087386 PCT/EP03/04010

	_	_
7	7	

Ile Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ser Lys Lys Ala Met
405 410 415

- Val Thr Phe Leu Asn Leu Ile Ala Gly Glu Pro Asp Ile Ser Arg Val 420 425 430
- Pro Ile Met Ile Asp Ser Ser Lys Trp Asp Val Ile Glu Ala Gly Leu 435 440 445
- Lys Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Met Lys Glu 450 455 460
- Gly Val Glu Gln Phe Ile His His Ala Lys Leu Cys Lys Arg Tyr Gly
 465 470 475 480
- Ala Ala Val Val Met Ala Phe Asp Glu Ala Gly Gln Ala Asp Thr 485 490 495
- Glu Ala Arg Lys Lys Glu Ile Cys Lys Arg Ser Tyr Asp Ile Leu Val 500 505 510
- Asn Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile 515 520 525
- Phe Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Val Asp 530 535 540
- Phe Ile Asn Ala Cys Ala Tyr Ile Arg Asp Glu Leu Pro Tyr Ala Leu 545 550 555 560
- Ser Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn 565 570 575
- Pro Val Arg Glu Ala Ile His Ser Val Phe Leu Leu Tyr Ala Ile Arg 580 585 590
- Ala Gly Leu Thr Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr 595 600 605
- Asp Gln Ile Pro Gln Glu Leu Arg Asp Ala Val Glu Asp Val Ile Leu 610 620
- Asn Arg Thr Pro Glu Gly Thr Asp Ala Leu Leu Ala Ile Ala Asp Lys 625 630 635 640
- Tyr Lys Gly Asp Gly Ser Val Lys Glu Ala Glu Thr Glu Glu Trp Arg
 645 650 655
- Gly Trp Asp Val Asn Lys Arg Leu Glu His Ala Leu Val Lys Gly Ile
 660 665 670
- Thr Thr His Ile Val Glu Asp Thr Glu Glu Ser Arg Gln Ser Phe Ala 675 680 685
- Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Ser Gly Met Asn Ile
 690 695 700
- Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Val 705 710 715 720
- Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile Pro Phe 725 730 735

- Ile Glu Leu Glu Lys Gly Asp Lys Pro Glu Ala Lys Gly Lys Ile Leu 740 745 750
- Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val 755 760 765
- Gly Val Val Leu Gly Cys Asn Gly Tyr Asp Ile Val Asp Leu Gly Val
 770 780
- Met Val Pro Ala Glu Lys Ile Leu Gln Val Ala Lys Glu Gln Lys Cys 785 790 795 800
- Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met 805 810 815
- Val His Val Ala Arg Glu Met Gln Arg Gln Asp Phe His Leu Pro Leu 820 825 830
- Met Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile 835 840 845
- Glu Pro Lys Tyr Ser Asn Asp Ala Val Val Tyr Val Thr Asp Ala Ser 850 855 860
- Arg Ala Val Gly Val Ala Thr Gln Leu Leu Ser Lys Glu Leu Lys Ala 865 870 875 880
- Gly Phe Val Gln Lys Thr Arg Glu Glu Tyr Ile Asp Val Arg Glu Arg 885 890 895
- Thr Ala Asn Arg Ser Ala Arg Thr Glu Arg Leu Ser Tyr Ala Ala Ala 900 905 910
- Ile Ala Lys Lys Pro Gln Phe Asp Trp Ala Thr Tyr Thr Pro Val Lys 915 920 925
- Pro Thr Phe Thr Gly Thr Arg Val Leu Asp Asn Ile Asp Leu Asn Val 930 935 940
- Leu Ala Glu Tyr Ile Asp Trp Thr Pro Phe Phe Ile Ser Trp Asp Leu 945 950 955 960
- Ala Gly Lys Phe Pro Arg Ile Leu Glu Asp Glu Val Val Gly Glu Ala 965 970 975
- Ala Thr Ala Leu Tyr Lys Asp Ala Arg Glu Met Leu Thr Lys Leu Ile 980 985 990
- Asp Glu Lys Leu Ile Ser Ala Arg Ala Val Phe Gly Phe Trp Pro Ala 995 1000 1005
- Asn Gln Val His Asp Asp Asp Ile Glu Leu Tyr Gly Asp Asp Gly Lys 1010 1015 1020
- Pro Met Ala Arg Leu His His Leu Arg Gln Gln Ile Ile Lys Thr Asp 1025 1030 1035 1040
- Gly Lys Pro Asn Phe Ser Leu Ala Asp Phe Val Ala Pro Lys Asp Ser 1045 1050 1055
- Glu Val Thr Asp Tyr Val Gly Gly Phe Ile Thr Thr Ala Gly Ile Gly

1060

1065

1070

Ala Glu Glu Val Ala Lys Ala Tyr Gln Asp Ala Gly Asp Asp Tyr Asn 1075 1080 1085

Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Cys Ala 1090 1095 1100

Glu Trp Leu His Gln Gln Val Arg Lys Glu His Trp Gly Tyr Ala Lys 1105 1110 1115 1120

Asp Glu Ala Leu Asp Asn Glu Ala Leu Ile Lys Glu Gln Tyr Ser Gly
1125 1130 1135

Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys

Ala Thr Leu Phe Ala Leu Leu Asp Pro Glu Ala Gln Glu Met Arg Ala 1155 1160 1165

Gly Arg Ser Gly Val Phe Leu Thr Glu His Tyr Ala Met Phe Pro Ala 1170 1175 1180

Ala Ala Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe 1185 1190 1195 1200

Ala Val Gly Lys Val Asp Lys Asp Gln Val Gln Ser Tyr Thr Ser Arg 1205 1210 1215

Lys Gly Gln Glu Leu Ser Leu Thr Glu Arg Trp Leu Ala Pro Asn Leu 1220 1225 1230

Gly Tyr Asp Asn 1235

<210> 27

<211> 3705

<212> DNA

<213> Pseudomonas aeruginosa

<220>

<221> CDS

<222> (1)..(3702)

<223> RPA01772

<400> 27

atg tcc agc ccg ctc acc gat cgc agc gcc cgc ctg caa gcc ctc cag

Met Ser Ser Pro Leu Thr Asp Arg Ser Ala Arg Leu Gln Ala Leu Gln

10 15

cac gcc ctc agg gaa cgt atc ctg atc ctc gat ggc ggc atg ggc acc
His Ala Leu Arg Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr
20 25 30

atg atc cag agc tac aag ctg gaa gag gcc gac tac cgc ggc gag cgc 144
Met Ile Gln Ser Tyr Lys Leu Glu Glu Ala Asp Tyr Arg Gly Glu Arg
35 40 45

ttc gcc gac tgg ccg agc gac gtg aaa ggc aac aac gac ctc ttg ctg
Phe Ala Asp Trp Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu
50
60

Let	u Se	r Ar	g Pr	g ga	p Vai	l Ile	cag Glr	g gco n Ala	ate Ile	e gag e Gli 7!	ı Lya	g gcc Ala	tac Tyi	c Lev	80 Bac	240
gc: Ala	gge Gly	gc. Ala	c ga a As	c ato p Ilo 8	e Lei	gag 1 Glu	acc Thr	aac Asr	acc Thi	: Phe	e aac e Asn	gco Ala	aco Thi	cag Glr 95	gtg Val	288
tco Sez	c cag	g gce n Ala	c gad a Asj 100	р Туз	ggc Gly	atg Met	cag Gln	tcg Ser 105	Lev	g gcc	tac Tyr	gaa Glu	Cto Lev 110	naA ı	gtc Val	336
GIV	i GIÀ	115	a Arg	g Lev	ı Ala	Arg	120	Val	Ala	Asp	Ala	Lys 125	Thr	Ala	gag Glu	384
Thr	130	Asp	D TAE	3 Pro	Arg	135	Val	Ala	Gly	Val	Leu 140	Gly	Pro	Thr	agc Ser	432
Arg 145	inr	СУЕ	ser	: Ile	150	Pro	Asp	Val	Asn	Asn 155	Pro	Gly	Tyr	Arg	aac Asn 160	480
val	inr	Pne	: Авр	165	Leu	Val	Glu	Asn	Tyr 170	Val	gag Glu	Ala	Thr	Arg 175	Gly	528
Leu	116	Glu	180	Gly	Ala	qaA	Leu	Ile 185	Leu	Ile	gag Glu	Thr	Ile 190	Phe	Asp	576
Inr	rea	195	Ala	Lys	Ala	Ala	Ile 200	Phe	Ala	Val	cag Gln	Gly 205	Val	Phe	Glu	624
JIU	210	GIÀ		GIA	Leu	Pro 215	Ile	Met	Ile	Ser	gga Gly 220	Thr	Ile	Thr	Asp	672
225	ser	GIÀ	arg	Tnr	230	Ser	Gly	Gln	Thr	Thr 235	gag Glu	Ala	Phe	Trp	Asn 240	720
,61	Val	AIG	uis	245	Arg	Pro	116	ser	Val 250	Gly	ctg Leu	Asn	Сув	Ala 255	Leu	768
ыл	AIA	гув	260	Leu	Arg	Pro	Tyr	11e 265	Glu	Glu	ctg Leu	Ser	Thr 270	Lув	Ala	816
rsb	Inr	275	vai	ser	Ala	His	Pro 280	Asn	Ala	Gly	ctg Leu	Pro 285	Asn	Ala	Phe	864
тy	290	lyr	Авр	GIU	ser	Pro 295	Ala	Glu	Met	Ala	gtg Val 300	Val	Val	Glu	Glu	912
CC	gcc	gcc	gcc	ggc	ttc	ctc	aat	atc	gtc	9gc	ggc	tgc	tgc	ggc	acc	960

_,										117						
Phe 305	Ala	Ala	a Ala	a Gly	7 Phe 310	e Leu	ı Ası	ı Ile	· Va]	315		Cys	Сув	Gly	Thr 320	
acc Thr	Pro) Ala	g cac His	325	e Glu	g gcg	ato Ile	gcc Ala	aag Lys 330	. Ala	gtg Val	gcc Ala	aag Lys	Tyr 335	ccg Pro	1008
ccg Pro	cgg Arg	gco Ala	ato Ile 340	Pro	gag Glu	ı att	ccc Pro	cgg Arg 345	Ala	tgt Cys	cgc Arg	ctg Leu	tcc Ser 350	Gly	ctg Leu	1056
gag Glu	ccg Pro	Phe 355	Thr	ato	gac Asp	cgc Arg	e ago Ser 360	Ser	ctg Leu	ttc Phe	gtc Val	aac Asn 365	Val	ggc Gly	gag Glu	1104
cgc Arg	acc Thr 370	Asn	atc Ile	Thr	ggt	tcg Ser 375	Ala	aag Lys	ttc Phe	gcc	cgg Arg 380	ctg Leu	atc Ile	cgc Arg	gag Glu	1152
gaa Glu 385	aac Asn	tac Tyr	gcg	gaa Glu	gct Ala 390	ctc Leu	gag Glu	gtc Val	gcc	cag Gln 395	cag Gln	cag Gln	gtg Val	gaa Glu	gcc Ala 400	1200
ggc	gcc Ala	cag Gln	gtg Val	atc Ile 405	gac Asp	atc Ile	aac Asn	atg Met	gac Asp 410	Glu	ggc Gly	atg Met	ctg Leu	gac Asp 415	tcg Ser	1248
Lys	Ala	Ala	Met 420	Val	Thr	Phe	Leu	Asn 425	Leu	Ile	gcc Ala	Ser	Glu 430	Pro	Asp	1296
atc Ile	tcg Ser	cgc Arg 435	Val	ccg Pro	atc Ile	atg Met	atc Ile 440	gac Asp	tcc Ser	tcc Ser	aag Lys	tgg Trp 445	gaa Glu	gtg Val	atc Ile	1344
Glu	gcc Ala 450	ggc	ctg Leu	aag Lys	tgc Cys	atc Ile 455	cag Gln	ggc Gly	aag Lys	ggc Gly	atc Ile 460	gtc Val	aac Asn	tcg Ser	atc Ile	1392
tcg Ser 465	atg Met	aag Lys	gaa Glu	ggc Gly	gtc Val 470	gag Glu	gcc Ala	ttc Phe	aag Lys	cac His 475	cat His	gcc Ala	cgc Arg	ctg Leu	tgc Cys 480	1440
aag Lys	cgc Arg	tac Tyr	ggc Gly	gcc Ala 485	gcg Ala	gtg Val	gtg Val	gtg Val	atg Met 490	gcc Ala	ttc Phe	gac Asp	gag Glu	gac Asp 495	ggc Gly	1488
cag Gln	gcc Ala	gac Asp	acc Thr 500	cag Gln	gcg Ala	cgc Arg	aag Lys	gaa Glu 505	gaa Glu	atc Ile	tgc Cys	aag Lys	cgc Arg 510	tcc Ser	tac Tyr	1536
gac Asp	TIE	ctg Leu 515	gtc Val	gac Asp	gaa Glu	Val	ggc Gly 520	ttc Phe	cca Pro	ccg Pro	Glu	gac Asp 525	atc Ile	atc Ile	ttc Phe	1584
gat (gcg Ala 530	aac Asn	atc Ile	ttc Phe	Ala	atc Ile 535	gcc Ala	acc Thr	ggc	atc Ile	gag Glu 540	gaa Glu	cac His	aac Asn	aac Asn	1632
tac q Tyr 1 545	gcg Ala	gtc Val	gat Asp	Phe	atc Ile 550	aac Asn	gcc Ala	tgc Cya	gcc Ala	tac Tyr 555	atc Ile	cgc Arg	gac Asp	aac Asn	ctc Leu 560	1680

Pro	tac Tyr	gcc Ala	ctg Leu	ago Ser 565	Ser	ggc	G1y	gtg Val	tcc Ser 570	Asn	gtg Val	tcc Ser	ttc Phe	tcg Ser 575	ttc Phe	· 1728
Arg	ggc	aac Asn	aac Asn 580	Pro	gta Val	cgc Arg	gag Glu	gcg Ala 585	Ile	cac His	tcg Ser	gtg Val	ttc Phe 590	ctc Leu	tac Tyr	1776
tac Tyr	gcg Ala	atc Ile 595	Arg	aac Asn	ggc	ctg Leu	acc Thr 600	atg Met	Gly	atc Ile	gtc Val	aac Asn 605	gcç Ala	ggc	cag Gln	1824
ctg Leu	gaa Glu 610	atc Ile	tac Tyr	gac Asp	gag Glu	att Ile 615	ccg Pro	aaa Lys	gcg Ala	ctg Leu	cgc Arg 620	gac Asp	cgg Arg	gtc Val	gag Glu	1872
gac Asp 625	Val	gtg Val	ctc Leu	aac Asn	cgc Arg 630	acg Thr	ccc Pro	gag Glu	gcc Ala	acc Thr 635	gag Glu	gcc Ala	ctg Leu	ctg Leu	gcg Ala 640	1920
atc Ile	gcc Ala	gac Asp	gac Asp	tac Tyr 645	aag Lys	ggc Gly	ggc Gly	ggc	gcg Ala 650	gtc Val	aag Lys	gag Glu	gcc Ala	gag Glu 655	gac Asp	1968
Glu	Glu	Trp	Arg 660	Ser	Tyr	agc Ser	Val	Glu 665	Lys	Arg	Leu	Glu	His 670	Ala	Leu	2016
Val	Lys	Gly 675	Ile	Thr	Thr	tgg Trp	Ile 680	Val	Glu	Ąsp	Thr	Glu 685	Glu	Сув	Arg	2064
Gln	Gln 690	Сув	Ala	Arg	Pro	atc Ile 695	Glu	Val	Ile	Glu	Gly 700	Pro	Leu	Met	Ser	2112
705	Met	Asn	Val	Val	Gly 710) Asp	Leu	Phe	Gly	Ala 715	Gly	Lys	Met	Phe	Leu 720	2160
Pro	Gln	Val	Val	Lys 725	Ser	gcg Ala	Arg	Val	Met 730	Lys	Gln	Ala	Val	Ala 735	His	2208
Leu	lle	Pro	740	He	Glu	gcg Ala	Glu	Lys 745	Gly	qaA	Lys	Pro	Glu 750	Ala	Lys	2256
ggc Gly	aag Lys	atc Ile 755	ctg Leu	atg Met	gcc Ala	acg Thr	gtg Val 760	aag Lys	ggc Gly	gac Asp	gtg Val	cac His 765	gac Asp	atc Ile	ggc Gly	2304
aag Lys	aac Asn 770	atc Ile	gtc Val	ggc Gly	gtg Val	gtg Val 775	ctc Leu	ggc Gly	tgc Cys	aac Asn	ggc Gly 780	tat Tyr	gac Asp	gtg Val	gtc Val	2352
gac Asp 785	ctc Leu	ggc Gly	gtg Val	Met	gtg Val 790	ccg Pro	gcg Ala	gag Glu	Lys	atc Ile 795	ctg Leu	cag Gln	acc Thr	gcc Ala	atc Ile 800	2400
gcc	gag	aaa	tgc ·	gac	atc	atc	ggc	ctg	tct	ggc	ctg	atc	acg	ccg	tcg	2448

										119 [°]						
Ala	Gli	ı Lyı	з Сув	805	lle i	: Ile	Gly	/ Lev	Ser 810	Gly	Leu	ılle	Thi	Pro 815		
ctg Leu	gac Asp	gag Glu	atg 1 Met 820	. Val	cac His	gto Val	gcc Ala	aag Lys 825	Glu	atg Met	cag Gln	cgg Arg	Cag Glr 830	Asn	ttc Phe	2496
cag Gln	Lev	9 ccc 1 Pro 835	g ctg Leu	atg Met	ato Ile	ggc	ggc Gly 840	Ala	act Thr	acc Thr	tcg Ser	aag Lys 845	Ala	cat His	acc Thr	2544
gcg Ala	gtg Val 850	. Lys	atc Ile	gat Asp	ccg Pro	Gln 855	Tyr	agc Ser	aac Asn	дас Авр	gcg Ala 860	Val	gtc Val	tac Tyr	gtc Val	2592
acc Thr 865	gac	gcc Ala	tcg Ser	cgc Arg	gcg Ala 870	Val	ggc Gly	gtg Val	gcc Ala	acc Thr 875	agc Ser	ctg Leu	ctg Leu	tcc Ser	aag Lys 880	2640
gag Glu	ctg Leu	aag Lys	gcc Ala	gac Asp 885	tac Tyr	gtg Val	gcc Ala	cgc Arg	acc Thr 890	Arg	gcc Ala	gac Asp	tac Tyr	gcg Ala 895	gtg Val	2688
gtc Val	cgc Arg	gaa Glu	cgc Arg 900	acg Thr	gcc Ala	aac Asn	cgc Arg	agc Ser 905	gcc Ala	cgc Arg	acc Thr	gag Glu	cgg Arg 910	ctg Leu	agc Ser	2736
tac Tyr	gaa Glu	cag Gln 915	gcg Ala	atc Ile	gcc Ala	aac Asn	aag Lys 920	ccg Pro	gcg Ala	ttc Phe	gac Asp	tgg Trp 925	gcc Ala	ggc	tac Tyr	2784
cag Gln	gcg Ala 930	ccg Pro	acg Thr	cct Pro	tcc Ser	ttc Phe 935	acc Thr	ggc Gly	gtc Val	agg Arg	gtg Val 940	ctc Leu	gac Asp	gag Glu	atc Ile	2832
gac Asp 945	ctc Leu	gcg Ala	gtg Val	ctc Leu	gcc Ala 950	gag Glu	tac Tyr	atc Ile	gac Asp	tgg Trp 955	acg Thr	ccg Pro	ttc Phe	ttc Phe	att Ile 960	2880
tcc Ser	tgg Trp	gac Asp	ctg Leu	gcc Ala 965	ggc Gly	aag Lys	tac Tyr	ccg Pro	cgc Arg 970	atc Ile	ctc Leu	acc Thr	gac Asp	gag Glu 975	gtg Val	2928
gtc Val	ggc Gly	gag Glu	gcc Ala 980	gcc Ala	acc Thr	tcg Ser	ttg Leu	ttc Phe 985	aac Asn	gac Asp	gcc Ala	cag Gln	gcg Ala 990	atg Met	ctg Leu	2976
aag Lys	aag Lys	ctg Leu 995	atc Ile	gac Asp	gag Glu	Lys	ctg Leu 000	atc Ile	aag Lys	gcc Ala	Arg	gcg Ala 005	gtg Val	ttc Phe	ggc Gly	3024
РЛЕ	tgg Trp 010	ccg Pro	gcc Ala	aac Asn	Gln	gtc Val 015	gag Glu	cac His	gac Asp	Asp	ctg Leu 020	gag Glu	gtc Val	tac Tyr	ggc Gly	3072
gcc Ala 1025	gat Asp	ggc Gly	gag Glu	Thr	ctc Leu 1030	Ala	acc Thr	ctg Leu	cac His	cac His 1035	Leu	cgg Arg	cag Gln	cag Gln	acg Thr 1040	3120
atc a	aag Lys	ccg Pro	gac Asp 1	ggc Gly 045	aag Lys	ccg Pro	aac Asn	Leu	tcg Ser 050	ctg Leu	gcc Ala	gat Asp	Phe	gtc Val .055	gcg Ala	3168

Pro	aag Lys	gaa Glu	agc Ser 1060	Gly	gtg Val	cgc	qaA	tac Tyr 1065	atc Ile	ggc Gly	ggc Gly	Phe	atc Ile 1070	Thr	acc Thr	3216
gcc Ala	GIA	atc Ile 1075	Gly	gcc Ala	gag Glu	Glu	gtg Val 1080	gcc Ala	aag Lys	gcg Ala	Tyr	gaa Glu 1085	gcc Ala	aag Lys	ggc Gly	3264
Asp	gac Asp 1090	Tyr	aac Asn	agc Ser	Ile	atg Met 1095	gtc Val	aag Lys	gcg	ctc Leu	gcc Ala 1100	gac Asp	cgc Arg	ctc Leu	gcc Ala	3312
110	A1a	Сув	Ala	Glu	Trp 1110	Leu)	His	Glu	Arg	gtg Val 111!	Arg	Lys	Glu	Tyr	Trp 1120	3360
ggc	tac Tyr	gcc Ala	Arg	gac Asp 1125	gaa Glu	cac His	ctc Leu	Asp	aac Asn 1130	gag Glu	gcc Ala	ttg Leu	Ile	aag Lys 1135	gag Glu	3408
caa Gln	tac Tyr	val	ggc Gly 1140	atc Ile	cgc Arg	ccg Pro	Ala	ccg Pro l145	ggc Gly	tac Tyr	ccg Pro	Ala	tgc Сув 1150	ccc Pro	gac Asp	3456
cat His	Thr	gag Glu 1155	aaa Lys	ggc	act Thr	Leu	ttc Phe 160	gaa Glu	ctg Leu	ctc Leu	Asp	ccg Pro 165	cag Gln	ggc	ctg Leu	3504
Ser	ggc Gly .170	gtc Val	agc Ser	ctg Leu	Thr	gag Glu 175	cac His	tac Tyr	gcg Ala	atg Met 1	ttc Phe 180	ccg Pro	gcc Ala	gcg Ala	gcg Ala	3552
gtc Val 1185	ser	ggt Gly	tgg Trp	tat Tyr	ttc Phe 1190	Ala	cac His	ccg Pro	cag Gln	gcg Ala 1195	Gln	tac Tyr	ttc Phe	gcg Ala	gtc Val 1200	3600
ggc Gly	aag Lys	atc Ile	Asp	aag Lys 205	gac Asp	cag Gln	gtg Val	Glu	cgc Arg 210	tac Tyr	agc Ser	cag Gln	Arg	aag Lys 215	ggc Gly	3648
cag Gln	gaa Glu	Ala	agc Ser 220	gtc Val	agc Ser	gag Glu	Arg	tgg Trp 225	ctg Leu	gcg Ala	ccg Pro	Asn	ctt Leu 230	ggc Gly	tac Tyr	3696
gat Asp	-	_														3705
<210 <211 <212 <213	> 12 > PR	34 T	mona	в ае	rugi	nosa										

Met Ser Ser Pro Leu Thr Asp Arg Ser Ala Arg Leu Gln Ala Leu Gln
1 5 10 15

His Ala Leu Arg Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr
20 25 30

<400> 28

WO 03/087386 PCT/EP03/04010

4	~4
1	71

Met Ile Gln Ser Tyr Lys Leu Glu Glu Ala Asp Tyr Arg Gly Glu Arg
35 40 45

Phe Ala Asp Trp Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu Leu 50 55 60

Leu Ser Arg Pro Asp Val Ile Gln Ala Ile Glu Lys Ala Tyr Leu Asp
65 70 75 80

Ala Gly Ala Asp Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Val 85 90 95

Ser Gln Ala Asp Tyr Gly Met Gln Ser Leu Ala Tyr Glu Leu Asn Val 100 105 110

Glu Gly Ala Arg Leu Ala Arg Gln Val Ala Asp Ala Lys Thr Ala Glu 115 120 125

Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Val Leu Gly Pro Thr Ser 130 140

Arg Thr Cys Ser Ile Ser Pro Asp Val Asn Asn Pro Gly Tyr Arg Asn 145 150 155 160

Val Thr Phe Asp Glu Leu Val Glu Asn Tyr Val Glu Ala Thr Arg Gly
165 170 175

Leu Ile Glu Gly Gly Ala Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp 180 185 190

Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Val Gln Gly Val Phe Glu 195 200 205

Glu Leu Gly Val Glu Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp 210 215 220

Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Thr Glu Ala Phe Trp Asn 225 230 235 240

Ser Val Arg His Ala Arg Pro Ile Ser Val Gly Leu Asn Cys Ala Leu 245 250 255

Gly Ala Lys Glu Leu Arg Pro Tyr Ile Glu Glu Leu Ser Thr Lys Ala 260 265 270

Asp Thr His Val Ser Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe 275 280 285

Gly Glu Tyr Asp Glu Ser Pro Ala Glu Met Ala Val Val Glu Glu 290 295 300

Phe Ala Ala Ala Gly Phe Leu Asn Ile Val Gly Gly Cys Cys Gly Thr 305 310 315 320

Thr Pro Ala His Ile Glu Ala Ile Ala Lys Ala Val Ala Lys Tyr Pro 325 330 335

Pro Arg Ala Ile Pro Glu Ile Pro Arg Ala Cys Arg Leu Ser Gly Leu 340 345 350

Glu Pro Phe Thr Ile Asp Arg Ser Ser Leu Phe Val Asn Val Gly Glu 355 360 365

Arg	370		lle	Thi	Gly	Ser 375		Lys	Phe	Ala	Arg 380		Ile	. Arg	Glu
Glu 385	' Asn	Tyr	Ala	Glu	Ala 390		Glu	Val	Ala	Gln 395		Gln	Val	Glu	Ala 400
Gly	Ala	Gln	Val	1le 405	Asp	Ile	aeA	Met	Asp 410		Gly	Met	Leu	Asp 415	
Lys	Ala	Ala	Met 420		Thr	Phe	Leu	Asn 425		Ile	Ala	Ser	Glu 430		Asp
Ile	Ser	Arg 435	Val	Pro	Ile	Met	Ile 440		Ser	Ser	Гув	Trp 445		Val	Ile
Glu	Ala 450	Gly	Leu	Lys	Сув	Ile 455		Gly	Lys	Gly	Ile 460		Asn	Ser	Ile
Ser 465	Met	Lys	Glu	Gly	Val 470	Glu	Ala	Phe	Lys	His 475	His	Ala	Arg	Leu	Сув 480
Lys	Arg	Tyr	Gly	Ala 485	Ala	Val	Val	Val	Met 490		Phe	Asp		Авр 495	
Gln	Ala	Asp	Thr 500	Gln	Ala	Arg	Lys	Glu 505		Ile	Сув	Lys	Arg 510		Туг
		515			Glu		520					525			•
Asp	Ala 530	Asn	Ile	Phe	Ala	Ile 535	Ala	Thr	Gly	Ile	Glu 540	Glu	His	Asn	Asn
545					Ile 550					555			_		560
				565	Ser				570					575	
			580		Val			585					590		
		595			Gly		600					605		_	
	610				Glu	615					620				
Asp 625	Val	Val	Leu	Asn	Arg 630	Thr	Pro	Glu	Ala	Thr 635	Glu	Ala	Leu	Leu	Ala 640
Ile	Ala	Asp	Asp	Tyr 645	Lys	Gly	Gly	Gly	Ala 650	Val	Lys	Glu	Ala	Glu 655	Авр
Glu	Glu	Trp	Arg 660	Ser	Tyr	Ser	Val	Glu 665	Lys	Arg	Leu	Glu	His 670	Ala	Leu
Val	Lys	Gly 675	Ile	Thr	Thr	Trp	Ile 680	Val	Glu	Asp	Thr	Glu 685	Glu	Сув	Arg

Gln Gln Cys Ala Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Ser

695 700 Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile Pro Phe Ile Glu Ala Glu Lys Gly Asp Lys Pro Glu Ala Lys Gly Lys Ile Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly 760 Lys Asn Ile Val Gly Val Val Leu Gly Cys Asn Gly Tyr Asp Val Val 780 Asp Leu Gly Val Met Val Pro Ala Glu Lys Ile Leu Gln Thr Ala Ile 790 Ala Glu Lys Cys Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser 805 Leu Asp Glu Met Val His Val Ala Lys Glu Met Gln Arg Gln Asn Phe 825 Gln Leu Pro Leu Met Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Asp Pro Gln Tyr Ser Asn Asp Ala Val Val Tyr Val 850 855 Thr Asp Ala Ser Arg Ala Val Gly Val Ala Thr Ser Leu Leu Ser Lys Glu Leu Lys Ala Asp Tyr Val Ala Arg Thr Arg Ala Asp Tyr Ala Val 890 Val Arg Glu Arg Thr Ala Asn Arg Ser Ala Arg Thr Glu Arg Leu Ser 905 Tyr Glu Gln Ala Ile Ala Asn Lys Pro Ala Phe Asp Trp Ala Gly Tyr Gln Ala Pro Thr Pro Ser Phe Thr Gly Val Arg Val Leu Asp Glu Ile 935 Asp Leu Ala Val Leu Ala Glu Tyr Ile Asp Trp Thr Pro Phe Phe Ile Ser Trp Asp Leu Ala Gly Lys Tyr Pro Arg Ile Leu Thr Asp Glu Val 970

Val Gly Glu Ala Ala Thr Ser Leu Phe Asn Asp Ala Gln Ala Met Leu

Lys Lys Leu Ile Asp Glu Lys Leu Ile Lys Ala Arg Ala Val Phe Gly 995 1000 1005

Phe Trp Pro Ala Asn Gln Val Glu His Asp Asp Leu Glu Val Tyr Gly

1015

WO 03/087386 PCT/EP03/04010

4	04
1	24

Ala Asp Gly Glu Thr Leu Ala Thr Leu His His Leu Arg Gln Gln Thr 1025 1030 1035 1040

Ile Lys Pro Asp Gly Lys Pro Asn Leu Ser Leu Ala Asp Phe Val Ala 1045 1050 1055

Pro Lys Glu Ser Gly Val Arg Asp Tyr Ile Gly Gly Phe Ile Thr Thr 1060 1065 1070

Ala Gly Ile Gly Ala Glu Glu Val Ala Lys Ala Tyr Glu Ala Lys Gly
1075 1080 1085

Asp Asp Tyr Asn Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala 1090 1095 1100

Glu Ala Cys Ala Glu Trp Leu His Glu Arg Val Arg Lys Glu Tyr Trp 1105 1110 1115 1120

Gly Tyr Ala Arg Asp Glu His Leu Asp Asn Glu Ala Leu Ile Lys Glu 1125 1130 1135

Gln Tyr Val Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp 1140 1145 1150

His Thr Glu Lys Gly Thr Leu Phe Glu Leu Leu Asp Pro Gln Gly Leu 1155 1160 1165

Ser Gly Val Ser Leu Thr Glu His Tyr Ala Met Phe Pro Ala Ala Ala 1170 1175 1180

Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe Ala Val 1185 1190 1195 1200

Gly Lys Ile Asp Lys Asp Gln Val Glu Arg Tyr Ser Gln Arg Lys Gly 1205 1210 1215

Gln Glu Ala Ser Val Ser Glu Arg Trp Leu Ala Pro Asn Leu Gly Tyr 1220 1225 1230

Asp Asp

<210> 29

<211> 3714

<212> DNA

<213> Nitrosomas europeae

<220>

<221>. CDS

<222> (1) .. (3711)

<223> RNE01732

<400> 29

atg aca atg cat gaa cgt gct gat ttg ctg aaa cgg ttg ctt gcc gag 48
Met Thr Met His Glu Arg Ala Asp Leu Leu Lys Arg Leu Leu Ala Glu
1 5 10

cgt atc ctg atg ctc gac ggt gcc atg ggt acg atg atc cag agc tac 96
Arg Ile Leu Met Leu Asp Gly Ala Met Gly Thr Met Ile Gln Ser Tyr
20 25 30

										25						
aaa Lys	ctg Leu	acc Thr 35	Glu	tcg Ser	gat Asp	tat Tyr	cgg Arg 40	Gly	gaa Glu	cgt Arg	ttt Phe	gcc Ala 45	gat Asp	ttt Phe	ccg Pro	144
cat His	gat Asp 50	Leu	aaa Lys	ggc	aac Asn	aat Asn 55	Asp	ctg Leu	ctc Leu	tgc Cys	ctg Leu 60	acc Thr	aga Arg	ccg Pro	gaa Glu	192
gtc Val 65	atc Ile	cgc Arg	tcc Ser	att	cat His 70	cgt Arg	gct Ala	tac Tyr	ctc Leu	gaa Glu 75	gcc Ala	ggg Gly	tcg Ser	gat Asp	atc Ile 80	24(
atc Ile	gag Glu	acc Thr	aac Asn	acg Thr 85	ttc Phe	aac Asn	tcg Ser	aat Asn	gcg Ala 90	ccg Pro	tcg Ser	atg Met	gcg Ala	gac Asp 95	tac Tyr.	288
cac His	atg Met	cag Gln	gat Asp 100	ctg Leu	gtg Val	tat Tyr	gaa Glu	ctg Leu 105	aat Asn	gtg Val	gcg Ala	ggt Gly	gcg Ala 110	cgc Arg	ctg Leu .	336
gcg Ala	tgt Cys	gag Glu 115	gaa Glu	gcg Ala	cgg Arg	gca Ala	atg Met 120	gaa Glu	acg Thr	cag Gln	caa Gln	cct Pro 125	gac Asp	cgg Arg	ccc Pro	384
cgt Arg	ttc Phe 130	gtt Val	gcc Ala	ggt Gly	gtg Val	atc Ile 135	ggg Gly	cct Pro	acc Thr	acc Thr	aaa Lys 140	acg Thr	gct Ala	tca Ser	ctc Leu	432
tca Ser 145	ccg Pro	gat Asp	gtc Val	aat Asn	gat Asp 150	cct Pro	gga Gly	ttc Phe	cgg Arg	gcc Ala 155	att Ile	acc Thr	ttc Phe	gat Asp	gat Asp 160	480
ctg Leu	gtg Val	gaa Glu	agc Ser	tat Tyr 165	acc Thr	gag Glu	tcg Ser	gtg Val	cgc Arg 170	ggg Gly	ctg Leu	atc Ile	gac Asp	gga Gly 175	ggc Gly	528
gcg Ala	gat Asp	att Ile	ctg Leu 180	ctg Leu	gtc Val	gaa Glu	acc Thr	att Ile 185	ttt Phe	gac Asp	acc Thr	ttg Leu	aat Asn 190	gcc Ala	aaa Lys	576
gcc Ala	gca Ala	ttg Leu 195	ttt Phe	gcc Ala	atc Ile	gat Asp	cag Gln 200	tat Tyr	ttc Phe	gaa Glu	acg Thr	cat His 205	gga Gly	tta Leu	cgt Arg	624
ctg Leu	ccg Pro 210	gtg Val	atg Met	ata Ile	tcg Ser	gtc Val 215	acg Thr	att Ile	acc Thr	gat Asp	gct Ala 220	tcg Ser	gga Gly	cgt Arg	aat Asn	672
ctt Leu 225	tcc Ser	G1y 999	cag Gln	aca Thr	ccg Pro 230	gaa Glu	gct Ala	ttc Phe	tgg Trp	aat Asn 235	tcg Ser	gta Val	cgg Arg	cat His	gca Ala 240	720
cgt Arg	ccg Pro	ctt Leu	tcg Ser	gtg Val 245	gga Gly	atc Ile	aac Asn	tgc Cys	gcg Ala 250	ttg Leu	ggt Gly	gcg Ala	gag Glu	ttg Leu 255	atg Met	768
cgc Arg	ccc Pro	tac Tyr	gtg Val 260	gaa Glu	gag Glu	ttg Leu	Ser	aat Asn 265	gtg Val	gct Ala	gag Glu	gtt Val	ttc Phe 270	acc Thr	agc Ser	816
gcc Ala	cat His	ccc Pro	aat Asn	gcc Ala	ggc Gly	ttg Leu	cct Pro	aat Asn	ccc Pro	ttg Leu	gcg Ala	gaa Glu	acc Thr	ggt Glv	tat Tyr	864

275	280	285

gac gaa acg ccg gaa tat acc gcc cgt ctg atc aag gat ttt gcg caa hap Glu Thr Pro Glu Tyr Thr Ala Arg Leu Ile Lys Asp Phe Ala Gln 290 tcc ggg ttc gtc aac att gtc ggc ggc tgc tgt ggc act aca ccg aaa 320 310 315 315 315 316 316 315 316 316 316 316 316 316 316 316 316 316																	
cat atc gcg gcc att gca gaa act ctc ctg ttc gaa ac gtg ggt gac ctg act atc gas gac att ccc ctg ttc gta acc ggc gcc att fac gcg gcc att gcc ctg ttc gta acc ggc gct att fac ggc gcc gcc gtc att atc gas gac att ccc tcg gcc gcc gcc gcc gcc gcc gcc g	Asp	· 29	u Thi	r Pro	o Gli	и Туг	295	Ala	Arç	J Let	ı Ile	300))	Pho	e Ale	Gln	912
ctg ccc gat att cct aaa aaa ctg agg ctt tcc ggc ctc gag ccg ctc leu Pro Asp Ile Pro Lys Lys Leu Arg Leu Ser Gly Leu Glu Pro Leu 340 aat atc gat gaa cat tcc ctg ttc gta aac gtg ggt gaa cgt acc aat Ass Ile Asp Glu His Ser Leu Phe Val Asn Val Gly Glu Arg Thr Ass 355 gtc acc ggc tcc aag gca ttt gcc cgg ctg att ctc aat ggc ggt tat 1152 gtc acc ggc tcc aag gca ttt gcc cgg ctg att ctc aat ggc ggt tat 1737 gct gaa ggg ctg gtg atc gcg cgc agc cag gtg gaa acc ggc caa 1200 Ala Glu Gly Leu Val Ile Ala Arg Leu Ile Leu Ass Gly Gly Tyr 370 gct gaa ggg ctg gtg atc gcg cgc agc cag gtg gad aac ggc gca caa 1200 Ala Glu Gly Leu Val Ile Ala Arg Ser Gln Val Glu Ass Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Gly Gly Ala Gly Gly Gly Ala Gly Gly Gly Ala Ala Gly Ala Gly Gly Ala Gly Gly Gly Ala Ala Gly Arg Gly Gly Gly Ala Ala Gly Gly Ala Ala Gly Gly Gly Ala Gly Gly Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Gly Ala Gly Gly Gly Ala Gly Gly Gly Gly Ala Gly Gly Ala Gly	ser	: GT	g tto / Phe	gto Val	aac l Asr) Ile	e Val	ggc Gly	ggc	tgo Cys	Сув	: Gly	act Thi	aca Thi	c Pro	Lys	960
aat atc gat gac cat tcc ctg ttc gta aac gtg ggt gaa cgt acc aat Assn Ile Asp Glu His Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn 355 gtc acc ggc tcc aag gca ttt gcc cgg ctg att ctc aat ggc ggt tat 1152 gtc acc ggc tcc aag gca ttt gcc cgg ctg att ctc aat ggc ggt tat 1152 gt aac ggc tcc aag gca ttt gcc cgg ctg att ctc aat ggc ggt tat 1152 gct gaa ggg ctg gtg atc gcg agc cag gtg gag aac ggc gca caa Ala Glu Gly Leu Val Ile Ala Arg Ser Gln Val Glu Asn Gly Ala Glu 390 atc atc gat atc aac atg gat gaa gcg atg ctg gat tca cag aag gcg Ile Ile Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Gln Lys Ala 415 atg gtg acc ttt ctg aat ctg ctc gct gcc gaa ccg gat atc acg aag gcg gt leu Ala	cat His	ato Ile	gcg Ala	geo Ala	ITE	: Ala	gaa Glu	gcg Ala	gta Val	Arg	, Asp	ato Ile	cct Pro	ccc Pro	Arg	Pro	1008
gtc acc ggc tcc aag gca ttt gcc gg ctg att ctc aat ggc ggt tat 1152 gtc acc ggc tcc aag gca ttt gcc gg ctg att ctc aat ggc ggt tat 7370 gct gaa ggg ctg gtg atc gcg cgc agc cag gtg gag aac ggc gca caa 1200 Ala Glu Gly Leu Val Ile Ala Arg Ser Gln Val Glu Asn Gly Ala Gln 395 atc atc gat atc aac atg gat gaa gcg atg ctg gat tca cag aag gcg 11248 Ile Ile Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Gln Lys Ala 415 atg gtg acc ttt ctg aat ctg ctc gct gcc gaa ccg gat acc agc cgg atg gg gad acc gg cgg atg atg gtg gat cag gat acc acc and 415 atg gtg acc ttt ctg aat ctg ctc gct gcc gaa ccg gat acc agc cgg 1296 Met Val Thr Phe Leu Asn Leu Leu Ala Ala Glu Pro Asp Ile Ser Arg 420 ctg ccg atc atg ctc gat tcc agc aaa tgg tcg gtg atc gaa gcc gga 1344 Leu Pro Ile Met Leu Asp Ser Ser Lys Trp Ser Val Ile Glu Ala Gly 445 ctg aca tgt gtc cag ggt aag gcg gtc atc aat tcc atc agc ctc aag 1392 ctg aca tgt gtc cag ggt aag gcg gtc atc aat tcc atc agc ctc aag 1392 ctg aca ggt gaa gcg gag ttt tta cat cat gcc agg ctg gcg cgt cgt tat Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 470 gga ggt gaa gcg gtg att gtc atg gct tcc gac gaa acc ggg cag gcc gat 1440 ggg gcc gcg gtg att gtc atg gct tcc gac gaa acc ggg cag gcc gat 1488 Glu Gly Glu Ala Olu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 475 acc ttg cag cgc aag gtg gaa atc gca gcc gt tgt tac cat aca ctg Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 495 acc ttg cag cgc aag gtg gaa atc tgc acc gct tgt tac cat aca ctg 1536 Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	ctg Leu	Pro	gat Asp	TTE	PTO	aaa Lys	aaa Lys	ctg Leu	Arg	Leu	tcc Ser	ggc	cto Lev	ı Glu	Pro	g ctc Leu	1056
gct gaa ggg ctg gtg atc gcg cgc agc cag gtg gag aac ggc gca caa l200 atc atc atc gat atc asc atg gat ctg gat ctg aat ctg ctc gct gtg atc gag gcg gat atc atc agc cgg atc atc atc gcd atc atc gat atc atc atc gat atc atc atc atc atc atc atc atc atc a	aat Asn	ato	Asp	GIU	cat His	tcc Ser	ctg Leu	Phe	gta Val	aac Asn	gtg Val	ggt Gly	Glu	Arg	acc	aat Asn	1104
atc atc gat atc aac atg gat gaa gcg atg ctg gat tca cag aag gcg ll248 Ile Ile Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Gln Ly8 Ala 415 atg gtg acc ttt ctg aat ctg ctc gct gcc gaa ccg gat atc agc cgg ll296 Met Val Thr Phe Leu Asn Leu Leu Ala Ala Glu Pro Asp Ile Ser Arg 420 ctg ccg atc atg ctc gat tcc agc aaa tgg tcg gtg atc gaa gcc gga la40 Leu Pro Ile Met Leu Asp Ser Ser Ly8 Trp Ser Val Ile Glu Ala Gly 445 ctg aaa tgt gtc cag ggt aag gcg gtc atc aat tcc atc agc ctc agg leu Ly8 Cy8 Val Gln Gly Ly8 Ala Val Ile Asn Ser Ile Ser Leu Ly8 450 gaa ggt gaa gcg gag ttt tta cat cat gcc agg ctg gcg cgt cgt tat Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 470 ggg gcc gcg gtg att gtc atg gcd ttc gac gac gcg gad acc ggg gcc gat leu Ala Arg Arg Tyr 480 ggg gcc gcg gtg att gtc atg gcd ttc gac gac gcg ttg tac cat acc ctg Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 495 acc ttg cag cgc gat ttc cca ccc gag gat atc atc att ttc gac ccc aat lea Glu Gln Arg Ly8 Val Glu Ile Cy8 Thr Arg Cy8 Tyr His Thr Leu Soo Soo Soo Soo Soo Soo Soo Soo Soo So	gtc Val	Thr	GIA	tcc Ser	aag Lys	gca Ala	Phe	gcc Ala	cgg Arg	ctg Leu	att Ile	Leu	Asn	ggc	ggt	tat Tyr	1152
atg gtg acc ttt ctg aat ctg ctc gct gcc gaa ccg gat atc agc cgg level at atg ctg ctg aat ctg ctg ctg aat ctg ctg ctg ctg gtg atc gaa gcc gga late atg ctg ctg ctg aat tcc agc aat tcc agc aat ctg ctg ctg stg atc gaa gcc gga late leu atg ctg leu atg atg ctg ctg aat ctg ctg aat tcc agc aat ctg ctg aat tcc agc aat tcc atg ctg stg atc gaa gcc gga late leu atg ctg aat ctg aat tcc atg ctg aag gtg leu atg atg ctg ctg ctg atg leu atg ctg ctg ctg leu atg ctg ctg ctg ctg ctg ctg ctg ctg ctg c	Ala	gaa Glu	61 y 999	ctg Leu	gtg Val	He	gcg Ala	cgc Arg	agc Ser	cag Gln	Val	gag Glu	aac Asn	ggc	gca Ala	Gln	1200
ctg ccg atc atg ctc gat tcc agc aaa tgg tcg gtg atc gaa gcc gga late leu Leu Pro Ile Met Leu Asp Ser Ser Lys Trp Ser Val Ile Glu Ala Gly 445 ctg aaa tgt gtc cag ggt aag gcg gtc atc atc atc agc ctc aag laga gct lys Cys Val Gln Gly Lys Ala Val Ile Asn Ser Ile Ser Leu Lys 450 gaa ggt gaa gcg gag ttt tta cat cat gcc agg ctg gcg cgt cgt tat laga ggt Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 475 ggg gcc gcg gtg att gtc atg gct ttc gac gaa acc ggg cag gcc gat laga Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 495 acc ttg cag cgc aag gtg gaa atc tgc acg cgt tgt tac cat aca ctg cat gcg cgt tgt tac cat acc tgc acg cgt tgt tac cat acc ggg cag gcc gat laga cac ttg cag cgc gat atc leu Cys Tyr His Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu Soo Soo Soo Soo Soo Soo Soo Soo Soo So	116	116	Авр	TIE	405	Met	Asp	Glu	Ala	Met 410	Leu	Asp	Ser	Gln	Lys 415	Ala	1248
Ctg aaa tgt gtc cag ggt aag gcg gtc atc aat tcc atc agc ctc aag Leu Lys Cys Val Gln Gly Lys Ala Val Ile Asn Ser Ile Ser Leu Lys 450 gaa ggt gaa gcg gag ttt tta cat cat gcc agg ctg gcg cgt cgt tat Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 465 ggg gcc gcg gtg att gtc atg gct ttc gac gaa acc ggg cag gcc gat Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 485 acc ttg cag cgc aag gtg gaa atc tgc acg cgt tgt tac cat aca ctg Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	Mec	vai	Inr	420	reu	Asn	Leu	Leu	Ala 425	Ala	Glu	Pro	Ąap	Ile 430	Ser	Arg	
gaa ggt gaa gcg gag ttt tta cat cat gcc agg ctg gcg cgt cgt tat l440 Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 470 475 480 ggg gcc gcg gtg att gtc atg gct ttc gac gaa acc ggg cag gcc gat l488 Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 495 acc ttg cag cgc aag gtg gaa atc tgc acg cgt tgt tac cat aca ctg l536 Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 505 505 510 att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat l1884 Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	Deu	PIO	435	Met	ren	Asp	ser	Ser 440	Lys	Trp	Ser	Val	Ile 445	Glu	Ala	Gly	1344
ggg gcc gcg gtg att gtc atg gct ttc gac gaa acc ggg cag gcc gat 1488 Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 485 acc ttg cag cgc aag gtg gaa atc tgc acg cgt tgt tac cat aca ctg Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	ren	450	сув	vaI	GIN	GIÀ	Lys 455	Ala	Val	Ile	Asn	Ser 460	Ile	Ser	Leu	Lys	1392
acc ttg cag cgc aag gtg gaa atc tgc acg cgt tgt tac cat aca ctg Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	Giu	ggt Gly	gaa Glu	gcg Ala	gag Glu	Pne	tta Leu	cat His	cat His	gcc Ala	Arg	ctg Leu	gcg Ala	cgt Arg	cgt Arg	Tyr	1440
att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	G]A 333	gcc Ala	gcg Ala	gtg Val	TIE	gtc Val	atg Met	gct Ala	Phe	qaA	gaa Glu	acc Thr	gjy aaa	cag Gln	Ala	gat Asp	1488
The Glu Gin Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn	acc Thr	ttg Leu	cag Gln	Arg	aag Lys	gtg Val	gaa Glu	Ile	Сув	acg Thr	cgt Arg	tgt Cys	tac Tyr	His	aca Thr	ctg Leu	1536
	att (GIU	GIN	gcc Ala	gat Asp	ttc Phe	Pro 1	Pro (gag Glu	gat Asp	atc Ile	Ile	Phe	gac Asp	ccc Pro	aat Asn	1584

										127						
Ile	Phe 530	e Ala	att Ile	: gct : Ala	acg Thr	ggt Gly 535	' Ile	gaa Glu	gaa Glu	cac His	agt Ser 540	Asn	tat Tyr	gca Ala	gtg Val	1632
gat Asp 545	Phe	ato lle	gag Glu	g gcg i Ala	Thr 550	His	gtc Val	ato Ile	cgg Arg	caa Gln 555	Thr	ctg Leu	Pro	tat Tyr	gcc Ala 560	1680
aaa Lys	gtc Val	agc Ser	999 Gly	ggt Gly 565	Val	tcc Ser	aat Asn	gtt Val	Ser 570	Phe	tcg Ser	ttc Phe	cgg Arg	ggt Gly 575	aac Asn	1728
gaa Glu	Pro	atc Ile	cgc Arg 580	Glu	gcc	att Ile	cat His	acc Thr 585	Ala	ttc Phe	ctg Leu	tat Tyr	cac His 590	Ala	gtc Val	1776
aag Lys	gca Ala	ggc Gly 595	atg Met	acc Thr	atg Met	ggt Gly	atc Ile 600	Val	aac Asn	gca Ala	ggt Gly	cag Gln 605	Leu	999 Gly	gtt Val	1824
tat Tyr	tcc Ser 610	gac Asp	att Ile	ccg Pro	ccc Pro	gat Asp 615	ctg Leu	ctg Leu	gaa Glu	cat His	gtc Val 620	gag Glu	gat Asp	gta Val	ctg Leu	1872
ctg Leu 625	aac Asn	cgg Arg	cgg Arg	cct Pro	gat Asp 630	gca Ala	acc Thr	gaa Glu	cgt Arg	ctg Leu 635	gtg Val	gag Glu	ttt Phe	gcg Ala	gaa Glu 640	1920
cat His	ttc Phe	aag Lys	gga Gly	cag Gln 645	aaa Lys	aag Lys	gag Glu	cag Gln	atc Ile 650	gaa Glu	gat Asp	ctg Leu	tcc Ser	tgg Trp 655	cgt Arg	1968
gat Asp	gaa Glu	ccg Pro	gtg Val 660	cgg Arg	cag Gln	cgc Arg	ctg Leu	att Ile 665	cat His	gca Ala	ctg Leu	gtc Val	agg Arg 670	ggt Gly	atc Ile	2016
agc Ser	acc Thr	tac Tyr 675	atc Ile	gtc Val	gag Glu	gat Asp	acc Thr 680	gag Glu	ctc Leu	gtc Val	cgg Arg	cag Gln 685	gag Glu	atc Ile	gac Asp	2064
agc Ser	cag Gln 690	gga Gly	ggc Gly	aag Lys	ccg Pro	atc Ile 695	gag Glu	gtg Val	atc Ile	gaa Glu	ggc Gly 700	ccg Pro	ctc Leu	atg Met	gac Asp	2112
ggc Gly 705	atg Met	aat Asn	gta Val	gtg Val	999 Gly 710	gat Asp	ctg Leu	ttt Phe	ggc Gly	gca Ala 715	ggc Gly	aag Lys	atg Met	ttt Phe	ctg Leu 720	2160
cca Pro	cag Gln	gtg Val	gtc Val	aag Lys 725	tcg Ser	gca Ala	cgg Arg	gtg Val	atg Met 730	aag Lys	cag Gln	gcg Ala	gtt Val	gcc Ala 735	tat Tyr	2208
ctg Leu	ttg Leu	ccg Pro	tac Tyr 740	atc Ile	gag Glu	gca Ala	Glu	aaa Lys 745	aaa Lys	att Ile	tcc Ser	ggc Gly	gac Asp 750	agc Ser	aag Lys	2256
ccc Pro	aag Lys	ggc Gly 755	aag Lys	gtg Val	gtg Val	atc Ile	gct Ala 760	acc Thr	gtc Val	aaa Lys	ggg ggg	gat Asp 765	gtg Val	cat His	gat Asp	2304
att Ile	ggc Gly	aag Lys	aat Asn	atc Ile	gtt Val	tcc Ser	gtc Val	gtg Val	ttg Leu	cag Gln	tgt Cys	aat Asn	aac Asn	ttt Phe	gaa Glu	2352

1015

1020

128 775 780 gtc atc aac atg ggg gtg atg gtc ccc agt gca cag att ctg gaa aca 2400 Val Ile Asn Met Gly Val Met Val Pro Ser Ala Gln Ile Leu Glu Thr 790 gca cgc cgt gaa cag gtc gat atg atc ggt ctg tcc ggc ctg atc acc Ala Arg Arg Glu Gln Val Asp Met Ile Gly Leu Ser Gly Leu Ile Thr cct tcg ctg gaa gaa atg gcg cat gtt gcc cgg gaa atg gag cgt gaa Pro Ser Leu Glu Glu Met Ala His Val Ala Arg Glu Met Glu Arg Glu 825 caa ttc acc gtt ccg ctg ctg atc ggt ggc gcc acc act tcg cgg atg 2544 Gln Phe Thr Val Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Met cat acg gca gtc aaa atc gca ccc cat tac ggt ggg gtg acc gta tgg 2592 His Thr Ala Val Lys Ile Ala Pro His Tyr Gly Gly Val Thr Val Trp 850 855 860 gtg ccg gat gcc agc cgg gca gtc ggg gtg tgc agc aat ctg atg tca 2640 Val Pro Asp Ala Ser Arg Ala Val Gly Val Cys Ser Asn Leu Met Ser cag gat ctg cgt gat gac tat gtc cgg cag gtc aag gcc gag cag gag 2688 Gln Asp Leu Arg Asp Asp Tyr Val Arg Gln Val Lys Ala Glu Gln Glu 890 aag agc cgg gtg cag cac cgc aac aag aaa ggg cca tcc aag ctc ctc 2736 Lys Ser Arg Val Gln His Arg Asn Lys Lys Gly Pro Ser Lys Leu Leu 900 905 act ttc gag gaa gcc cgg gcc aac gca ctc aag acg gat tgg gct cgt Thr Phe Glu Glu Ala Arg Ala Asn Ala Leu Lys Thr Asp Trp Ala Arg 915 920 tat act cca cca gct ccg gat ttc ctg ggg ttg cgc acc ctc aac aac Tyr Thr Pro Pro Ala Pro Asp Phe Leu Gly Leu Arg Thr Leu Asn Asn 930 935 tat ccg ctg gaa aca ctg gtg ccg cac atc gac tgg aca cct ttc ttc Tyr Pro Leu Glu Thr Leu Val Pro His Ile Asp Trp Thr Pro Phe Phe 950 cag gca tgg gaa ctg cac ggg cgc tat cct gcc atc ctg cag gat gaa 2928 Gln Ala Trp Glu Leu His Gly Arg Tyr Pro Ala Ile Leu Gln Asp Glu ctc gtc ggg gaa gca gcc agc aat ctg ttt cgc gat gcc cag aat atg 2976 Leu Val Gly Glu Ala Ala Ser Asn Leu Phe Arg Asp Ala Gln Asn Met ctc aga aaa atc gtc gag caa aaa tgg ctc acc gcc aac gcc gtt atc 3024 Leu Arg Lys Ile Val Glu Gln Lys Trp Leu Thr Ala Asn Ala Val Ile 1000 ggc ctg ttc ccg gcc aat acc gtc aat gga gat gat atc gag att tat 3072 Gly Leu Phe Pro Ala Asn Thr Val Asn Gly Asp Asp Ile Glu Ile Tyr

			•	129	
1025	g ser Arg	Ser Gln	Val Ile Met	acc tgg cac acc to Thr Trp His Thr Le 1035	u Arg 1040
Gin Gin Thi	r Ala Lys 1045	Pro Ala	Gly Arg Pro 1050		a Asp 5
ttc att gcg Phe Ile Ala	g ccg cgt Pro Arg 1060	gaa acc g	gga ctg gac Gly Leu Asp 1065	gat acc atc ggt tt Asp Thr Ile Gly Le 1070	g ttt 3216 u Phe
Ala Val Ser 1075	Ala Gly	Phe Gly :	Ile Asp Glu 080	cgc ata cgc gct tt Arg Ile Arg Ala Ph 1085	e Glu
1090	Asp Asp	Tyr. Ser 1 1095	Ala Ile Ile	ctg aaa gca ctg gc Leu Lys Ala Leu Al 1100	a Āsp
1105	Glu Ala	Phe Ala C	3lu His Met	cat gca cgg gtg cg His Ala Arg Val Ar 1115	g Arg 1120
GIU Phe Trp	1125	Val Lys A	Asp Glu Ser 1130	ctg gac aat gaa ca Leu Asp Asn Glu Gl 113	n Leu 5
lie Asp Glu	caa tac Gln Tyr 1140	ctg gga a Leu Gly I	atc cgt cca [le Arg Pro 1145	gca cca ggt tat cc Ala Pro Gly Tyr Pro 1150	t gcc 3456 o Ala
tgc cct gat Cys Pro Asp 1155	cat acc His Thr	Glu Lys G	ggg cca ttg Sly Pro Leu .60	ttc gct ctg ctg ga. Phe Ala Leu Leu Gl 1165	a gcg 3504 u Ala
gaa aaa cgc Glu Lys Arg 1170	agc gga Ser Gly	atc gtc a Ile Val I 1175	ta acg gaa le Thr Glu	tca ttt gcc atg gt Ser Phe Ala Met Va 1180	g ccg 3552 l Pro
act gca gca Thr Ala Ala 1185	gta tcc Val Ser	ggc ttc t Gly Phe T 1190	at ctc tct yr Leu Ser	tac cct gaa tcc ago Tyr Pro Glu Ser Ser 1195	tat 3600 Tyr 1200
ttt gct gtt Phe Ala Val	gga aaa Gly Lys 1205	atc gga a Ile Gly L	aa gat cag ys Asp Gln 1210	gtc gag gat tat gca Val Glu Asp Tyr Ala 1215	Arg
Arg Lys Gly	tgg acg Trp Thr 220	ctg gaa g Leu Glu G	aa gca gaa lu Ala Glu 1225	agg tgg ctt gcg cct Arg Trp Leu Ala Pro 1230	gtc 3696 Val
ttg gcg tat Leu Ala Tyr 1235	gag cgt Glu Arg	taa			3714

<210> 30

<211> 1237

<212> PRT

<213> Nitrosomas europeae

- <400> 30
- Met Thr Met His Glu Arg Ala Asp Leu Leu Lys Arg Leu Leu Ala Glu 1 5 10 15
- Arg Ile Leu Met Leu Asp Gly Ala Met Gly Thr Met Ile Gln Ser Tyr
 20 25 30
- Lys Leu Thr Glu Ser Asp Tyr Arg Gly Glu Arg Phe Ala Asp Phe Pro 35 40 45
- His Asp Leu Lys Gly Asn Asn Asp Leu Leu Cys Leu Thr Arg Pro Glu
 50 55 60
- Val Ile Arg Ser Ile His Arg Ala Tyr Leu Glu Ala Gly Ser Asp Ile
 65 70 75 80
- Ile Glu Thr Asn Thr Phe Asn Ser Asn Ala Pro Ser Met Ala Asp Tyr
 85 90 95
- His Met Gln Asp Leu Val Tyr Glu Leu Asn Val Ala Gly Ala Arg Leu
 100 105 110
- Ala Cys Glu Glu Ala Arg Ala Met Glu Thr Gln Gln Pro Asp Arg Pro 115 120 125
- Arg Phe Val Ala Gly Val Ile Gly Pro Thr Thr Lys Thr Ala Ser Leu 130 135 140
- Ser Pro Asp Val Asn Asp Pro Gly Phe Arg Ala Ile Thr Phe Asp Asp 145 150 155 160
- Leu Val Glu Ser Tyr Thr Glu Ser Val Arg Gly Leu Ile Asp Gly Gly
 165 170 175
- Ala Asp Ile Leu Leu Val Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys
- Ala Ala Leu Phe Ala Ile Asp Gln Tyr Phe Glu Thr His Gly Leu Arg
- Leu Pro Val Met Ile Ser Val Thr Ile Thr Asp Ala Ser Gly Arg Asn 210 215 220
- Leu Ser Gly Gln Thr Pro Glu Ala Phe Trp Asn Ser Val Arg His Ala 225 230 240
- Arg Pro Leu Ser Val Gly Ile Asn Cys Ala Leu Gly Ala Glu Leu Met
 245 250 255
- Arg Pro Tyr Val Glu Glu Leu Ser Asn Val Ala Glu Val Phe Thr Ser 260 265 270
- Ala His Pro Asn Ala Gly Leu Pro Asn Pro Leu Ala Glu Thr Gly Tyr
 275 280 285
- Asp Glu Thr Pro Glu Tyr Thr Ala Arg Leu Ile Lys Asp Phe Ala Gln
 290 295 300
- Ser Gly Phe Val Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Lys 315 320
- His Ile Ala Ala Ile Ala Glu Ala Val Arg Asp Ile Pro Pro Arg Pro

- Leu Pro Asp Ile Pro Lys Lys Leu Arg Leu Ser Gly Leu Glu Pro Leu 340 345 350
- Asn Ile Asp Glu His Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn 355 360 365
- Val Thr Gly Ser Lys Ala Phe Ala Arg Leu Ile Leu Asn Gly Gly Tyr 370 380
- Ala Glu Gly Leu Val Ile Ala Arg Ser Gln Val Glu Asn Gly Ala Gln 385 390 395 400
- Ile Ile Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Gln Lys Ala
 405 410 415
- Met Val Thr Phe Leu Asn Leu Leu Ala Ala Glu Pro Asp Ile Ser Arg 420 425 430
- Leu Pro Ile Met Leu Asp Ser Ser Lys Trp Ser Val Ile Glu Ala Gly
 435 440 445
- Leu Lys Cys Val Gln Gly Lys Ala Val Ile Asn Ser Ile Ser Leu Lys 450 460
- Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 465 470 475 480
- Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 495
- Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 505 510
- Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn 515 520 525
- Ile Phe Ala Ile Ala Thr Gly Ile Glu Glu His Ser Asn Tyr Ala Val 530 540
- Asp Phe Ile Glu Ala Thr His Val Ile Arg Gln Thr Leu Pro Tyr Ala 545 550 555 560
- Lys Val Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn 565 570 575
- Glu Pro Ile Arg Glu Ala Ile His Thr Ala Phe Leu Tyr His Ala Val
- Lys Ala Gly Met Thr Met Gly Ile Val Asn Ala Gly Gln Leu Gly Val 595 605
- Tyr Ser Asp Ile Pro Pro Asp Leu Leu Glu His Val Glu Asp Val Leu 610 620
- Leu Asn Arg Arg Pro Asp Ala Thr Glu Arg Leu Val Glu Phe Ala Glu 625 630 635 640
- His Phe Lys Gly Gln Lys Lys Glu Gln Ile Glu Asp Leu Ser Trp Arg
 645 650 655

WO 03/087386 PCT/EP03/04010

132

Asp Glu Pro Val Arg Gln Arg Leu Ile His Ala Leu Val Arg Gly Ile
660 665 670

Ser Thr Tyr Ile Val Glu Asp Thr Glu Leu Val Arg Gln Glu Ile Asp 675 680 685

Ser Gln Gly Gly Lys Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp 690 695 700

Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu 705 710 715 720

Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr
725 730 735

Leu Leu Pro Tyr Ile Glu Ala Glu Lys Lys Ile Ser Gly Asp Ser Lys 740 745 750

Pro Lys Gly Lys Val Val Ile Ala Thr Val Lys Gly Asp Val His Asp 755 760 765

Ile Gly Lys Asn Ile Val Ser Val Val Leu Gln Cys Asn Asn Phe Glu 770 780

Val Ile Asn Met Gly Val Met Val Pro Ser Ala Gln Ile Leu Glu Thr 785 790 795 800

Ala Arg Arg Glu Gln Val Asp Met Ile Gly Leu Ser Gly Leu Ile Thr 805 810 815

Pro Ser Leu Glu Glu Met Ala His Val Ala Arg Glu Met Glu Arg Glu 820 825 830

Gln Phe Thr Val Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Met 835 840 845

His Thr Ala Val Lys Ile Ala Pro His Tyr Gly Gly Val Thr Val Trp 850 855 860

Val Pro Asp Ala Ser Arg Ala Val Gly Val Cys Ser Asn Leu Met Ser 865 870 875 880

Gln Asp Leu Arg Asp Asp Tyr Val Arg Gln Val Lys Ala Glu Gln Glu 885 890 895

Lys Ser Arg Val Gln His Arg Asn Lys Lys Gly Pro Ser Lys Leu Leu 900 905 910

Thr Phe Glu Glu Ala Arg Ala Asn Ala Leu Lys Thr Asp Trp Ala Arg 915 920 925

Tyr Thr Pro Pro Ala Pro Asp Phe Leu Gly Leu Arg Thr Leu Asn Asn 930 935 940

Tyr Pro Leu Glu Thr Leu Val Pro His Ile Asp Trp Thr Pro Phe Phe 945 950 955 960

Gln Ala Trp Glu Leu His Gly Arg Tyr Pro Ala Ile Leu Gln Asp Glu 965 970 975

Leu Val Gly Glu Ala Ala Ser Asn Leu Phe Arg Asp Ala Gln Asn Met 980 985 990

- Leu Arg Lys Ile Val Glu Gln Lys Trp Leu Thr Ala Asn Ala Val Ile 995 1000 1005
- Gly Leu Phe Pro Ala Asn Thr Val Asn Gly Asp Asp Ile Glu Ile Tyr 1010 1015 1020
- Ala Asp Arg Ser Arg Ser Gln Val Ile Met Thr Trp His Thr Leu Arg 1025 1030 1035 1040
- Gln Gln Thr Ala Lys Pro Ala Gly Arg Pro Asn Leu Ala Leu Ala Asp 1045 1050 1055
- Phe Ile Ala Pro Arg Glu Thr Gly Leu Asp Asp Thr Ile Gly Leu Phe 1060 1065 1070
- Ala Val Ser Ala Gly Phe Gly Ile Asp Glu Arg Ile Arg Ala Phe Glu 1075 1080 1085
- Ala Ala Asn Asp Asp Tyr Ser Ala Ile Ile Leu Lys Ala Leu Ala Asp 1090 1095 1100
- Arg Leu Ala Glu Ala Phe Ala Glu His Met His Ala Arg Val Arg Arg 1105 1110 1115 1120
- Glu Phe Trp Gly Tyr Val Lys Asp Glu Ser Leu Asp Asn Glu Gln Leu 1125 1130 1135
- Ile Asp Glu Gln Tyr Leu Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala 1140 1145 1150
- Cys Pro Asp His Thr Glu Lys Gly Pro Leu Phe Ala Leu Leu Glu Ala 1155 1160 1165
- Glu Lys Arg Ser Gly Ile Val Ile Thr Glu Ser Phe Ala Met Val Pro 1170 1175 1180
- Thr Ala Ala Val Ser Gly Phe Tyr Leu Ser Tyr Pro Glu Ser Ser Tyr 1185 1190 1195 1200
- Phe Ala Val Gly Lys Ile Gly Lys Asp Gln Val Glu Asp Tyr Ala Arg 1205 1210 1215
- Arg Lys Gly Trp Thr Leu Glu Glu Ala Glu Arg Trp Leu Ala Pro Val 1220 1225 1230

Leu Ala Tyr Glu Arg 1235

<210> 31

<211> 3774

<212> DNA

<213> Bordetella pertussis

<220>

<221> CDS

<222> (1) . . (3771)

<223> RBP00104

<220>

<221> unsure

			_							. 04						
			2		ces	of n	ind:	icate	any	/ nuc	cleo	tide				
<22	1> 1 2> 2		2		ces (of n	indi	lcat <i>e</i>	e anv	f fine	-len	ah i-	•			
<40	0 > 3	31														
gtg Val	Pro	tai	t cco) Ar	t ato	c cco e Pro	tto Phe	e Pro	ctg Leu 10	Ser	gcc Ala	tac Tyr	acc	cat His	ggc Gly	48
ggc	gag Glu	tto Phe	gto Val	LArg	caa g Gli	a cto n Lev	g gad 1 Asp	aag Lys 25	Arg	atc Ile	ctg Lev	g atc	cto Lev	ı Asp	ggt Gly	96
gcc Ala	atg Met	990 35	Thr	ato Met	; ato	cag Glm	g cgc Arg 40	Tyr	aag Lys	ctg Leu	ggo	gag Glu 45	Ala	gat Asp	ttc Phe	144
cgt Arg	ggc Gly 50	Glu	r cgc	tto Phe	gco Ala	gag Glu 55	His	cac His	aag Lys	gat Asp	ctc Leu 60	Lys	ggc	gac Yap	aac Asn	192
gaa Glu 65	ctg Leu	ctg Leu	tcg Ser	ntg Xaa	gtg Val 70	Arg	ccg Pro	gac Asp	gtg Val	atc Ile 75	Ala	gaa Glu	atc Ile	Cac	cgg Arg 80	240
cag Gln	tac Tyr	Ctc	gag Glu	gcc Ala 85	GIA	gcc Ala	gac Asp	gtg Val	atc Ile 90	gag Glu	acc Thr	nac Xaa	acc Thr	ttc Phe 95	ggc Gly	288
gcc Ala	acg Thr	tcg Ser	atc Ile 100	ATA	cag Gln	Gly	gat Asp	tac Tyr 105	gac Asp	ctg Leu	ccg Pro	gag Glu	ctg Leu 110	gcc Ala	tac Tyr	336
gag Glu	atg Met	aac Asn 115	ctg Leu	gag Glu	tcg Ser	gcc Ala	cgc Arg 120	ctg Leu	gcg Ala	cgc Arg	gcc Ala	gcc Ala 125	tgc Cys	gac Asp	gcc Ala	384
tac Tyr	agc Ser 130	acg Thr	ccc Pro	gag Glu	cat His	ccg Pro 135	cgc Arg	ttc Phe	gtg Val	gcc Ala	999 Gly 140	gcg Ala	ctg Leu	999 Gly	ccg Pro	432
cag Gln 145	ccc Pro	aag Lys	acc Thr	gcg Ala	tcc Ser 150	atc Ile	tcg Ser	ccc Pro	gac Asp	gtc Val 155	aac Asn	gac Asp	ccg Pro	999 Gly	gcg Ala 160	480
cgc Arg	aac Asn	gtc Val	acc Thr	ttc Phe 165	gac Asp	gag Glu	ctg Leu	cgc Arg	gcg Ala 170	gcc Ala	tat Tyr	gtc Val	gag Glu	cag Gln 175	ctc Leu	528
aat Asn	ggc Gly	ctg Leu	ctc Leu 180	yab	ggc Gly	ggc Gly	atc Ile	gac Asp 185	atc Ile	gtc Val	ctg Leu	atc Ile	gaa Glu 190	acc Thr	atc Ile	576
tc Phe	мвр	acg Thr 195	ctc Leu	aac Asn	gcc Ala	aag Lys	gcg Ala 200	gcc Ala	atc Ile	ttc Phe	gcc Ala	gtc Val 205	gag Glu	gaa Glu	gcg Ala	624
tc	qaq	qcq	cac	aac	ata	cac	cta	cca	ata	250	2++	+			_ • -	

WO 03/087386 PCT/EP03/04010

									1	35						
Phe	Glu 210	Ala	Arg	Gly	Val	Arg 215	Leu	Pro	Val	Met	Ile 220	Ser	Gly	Thr	Val	
acc Thr 225	gat Asp	gcg Ala	tcg Ser	ggc Gly	cgc Arg 230	atc Ile	ctg Leu	tcc Ser	ggc Gly	cag Gln 235	acc Thr	gtc Val	gag Glu	gcg Ala	ttc Phe 240	720
tgg Trp	aac Asn	tcg Ser	gtg Val	cgc Arg 245	cat	gcg Ala	cgg Arg	ccg Pro	gtc Val 250	acc Thr	atc Ile	ggc Gly	ctg Leu	aac Asn 255	tgc Cys	768
gcg Ala	ctg Leu	ggc	gcg Ala 260	gcg Ala	ctg Leu	atg Met	cgt Arg	ccg Pro 265	tat Tyr	gtg Val	gcc Ala	gag Glu	ctg Leu 270	tcc Ser	aag Lys	816
atc Ile	tgc Cys	gac Asp 275	acc Thr	tat Tyr	gtg Val	tgc Cys	gtc Val 280	tat Tyr	ccc Pro	aac Asn	gcc Ala	ggc Gly 285	ctg Leu	ccc Pro	aat Asn	864
ccc Pro	atg Met 290	gcc Ala	gag Glu	acg Thr	ggc Gly	ttt Phe 295	gac Asp	gaa Glu	acg Thr	ccg Pro	gcc Ala 300	gat Asp	acc Thr	tcg Ser	gcc Ala	912
ctg Leu 305	ctg Leu	gaa Glu	gag Glu	ttc Phe	gcc Ala 310	cag Gln	gcc Ala	elå aaa	ctg Leu	gtc Val 315	aac Asn	atg Met	gcc Ala	ggc Gly	ggc Gly 320	960
tgt Cys	tgc Cys	ggc Gly	acc Thr	acg Thr 325	ccc Pro	gag Glu	cac His	atc Ile	cgc Arg 330	gcc Ala	atc Ile	gcc Ala	ggc	aag Lys 335	gtg Val	1008
gcc Ala	gcg Ala	ctg Leu	acg Thr 340	ccg Pro	cgc Arg	gcg Ala	gtg Val	ccc Pro 345	gag Glu	gtg Val	ccg Pro	gtc Val	aag Lys 350	acc Thr	cgc Arg	1056
ctg Leu	tcg Ser	ggc Gly 355	ctg Leu	gag Glu	gcg Ala	ctc Leu	aac Asn 360	atc Ile	gac Asp	gac Asp	gag Glu	act Thr 365	ctg Leu	ttc Phe	gtc Val	1104
aac Asn	gtg Val 370	ggc Gly	gag Glu	cgc Arg	acc Thr	aac Asn 375	gtg Val	acg Thr	ggc Gly	agc Ser	aag Lys 380	atg Met	ttc Phe	gcc Ala	cgc Arg	1152
ctg Leu 385	gtc Val	cgc Arg	gag Glu	gag Glu	ааа L ув 390	tac Tyr	gac Asp	gag Glu	gcg Ala	ctg Leu 395	gcc Ala	gtg Val	gcg Ala	cgc Arg	cag Gln 400	1200
cag Gln	gtc Val	gag Glu	aac Asn	999 Gly 405	gcc Ala	cag Gln	atc Ile	atc Ile	gac Asp 410	gtc Val	aac Asn	atg Met	gac Asp	gag Glu 415	gcg Ala	1248
atg Met	ctg Leu	gac Asp	tcg Ser 420	gtg Val	gcc Ala	tgt Cys	atg Met	cac His 425	cgc Arg	ttc Phe	ctc Leu	aac Asn	ctg Leu 430	atc Ile	gcg Ala	1296
tcc Ser	gag Glu	ccc Pro 435	gac Asp	atc Ile	gcg Ala	Arg	gtg Val 440	ccg Pro	gtg Val	atg Met	atc Ile	gac Asp 445	agt Ser	tcc Ser	aag Lys	1344
Trp	gaa Glu 450	gtg Val	atc Ile	gag Glu	Thr	ggc Gly 455	ctg Leu	aag Lys	tgc Cys	gtg Val	cag Gln 460	ggc Gly	aag Lys	gcc Ala	gtg Val	1392

gtc Val 465	Asn	tcg Ser	atc Ile	tcc Ser	atg Met 470	Lys	gaa Glu	ggc Gly	gag Glu	gag Glu 475	ccg Pro	ttc Phe	cgc Arg	cat His	cat His 480	1440
gcg Ala	cgc Arg	ctg Leu	tgc Cys	cgc Arg 485	cgc Arg	tac Tyr	ggc	gcg Ala	gcc Ala 490	atg Met	gtg Val	gtc Val	atg Met	gcc Ala 495	ttc Phe	1488
gac Asp	gaa Glu	cag Gln	999 Gly 500	cag Gln	gcc Ala	Asp	tcg Ser	ctg Leu 505	gag Glu	cgc Arg	cgc Arg	aag Lys	gaa Glu 510	atc Ile	tgc Cys	1536
ggc	cgc Arg	gcc Ala 515	Tyr	cgt Arg	atc Ile	ctg Leu	gtc Val 520	gag Glu	gaa Glu	gag Glu	ggc	ttc Phe 525	ccg Pro	ccc Pro	gag Glu	1584
Asp	atc Ile 530	atc Ile	ttc Phe	gat Asp	ccc Pro	aac Asn 535	gtg Val	ttc Phe	gcg Ala	gtg Val	gcc Ala 540	Thr	ggc	atc Ile	gac Asp	1632
Glu 545	His	Asn	His	Tyr	Ala 550	Val	Asp	Phe	Ile	Glu 555	Gly	Ala	cgc Arg	Trp	Ile 560	1680
Arg	Ala	Asn	Leu	Pro 565	His	Ala	Arg	Ile	Ser 570	Gly	Gly	Ile		Asn 575	Val	1728
Ser	Phe	Ser	Phe 580	Arg	Gly	Asn	Glu	Pro 585	Met	Arg	Glu	Ala	atc Ile 590	His	Thr	1776
Val	Phe	Leu 595	Tyr	Tyr	Ala	Ile	Glu 600	Ala	Gly	Leu	Thr	Met 605	ggc	Ile	Val	1824
Asn	Ala 610	Gly	Gln	Leu	Gly	Val 615	Tyr	Ala	Asp	Leu	Ala 620	Pro	cac His	Leu	Arg	1872
gac Asp 625	ctg Leu	gtc Val	gag Glu	gac Asp	gtc Val 630	atc Ile	ctg Leu	gac Asp	cgc Arg	ccc Pro 635	gag Glu	ccg Pro	gtg Val	Gly	cgc Arg 640	1920
Ser	Asp	Ser	Ala	Asp 645	Glu	Arg	Ser	Pro	Thr 650	Glu	Arg	Leu	gtg Val	Gln 655	Phe	1968
gcc Ala	gag Glu	acc Thr	gtc Val 660	aag Lys	ggc Gly	tcg Ser	ggc Gly	gcg Ala 665	aag Lys	aag Lys	gaa Glu	gaa Glu	gac Asp 670	ctg Leu	acc Thr	2016
tgg Trp	cgc Arg	acc Thr 675	ggc Gly	tcg Ser	gtc Val	gag Glu	cag Gln 680	cgc Arg	ctg Leu	gcg Ala	cat His	gcc Ala 685	ctg Leu	gtg Val	cac His	2064
ggc Gly	atc Ile 690	acc Thr	acc Thr	ttc Phe	Ile	gtc Val 695	gag Glu	gac Asp	acc Thr	gag Glu	gaa Glu 700	gtg Val	cgc Arg	cag Gln	cag Gln	2112
gtc	gcc	gcg	cgc	ggc	999	cgc	acc	atc	gaa	gtg	atc	gaa	ggt	ccg	ctg	2160

										137						
Va] 709	l Al	a Al	a Ar	g Gl	y Gl	y Ar	g Th	r Ile	e Gli	715	1 116 5	e Glu	ı Gl	y Pro	720	
at <u>c</u> Met	: As	c gg p Gl	c at y Me	g aa t As: 72:	n Va.	g gto	c ggd l Gly	y As <u>r</u>	C Cto D Let 730	1 Phe	996 Gl	c gcg y Ala	999 Gly	c aaq y Ly: 73!	g atg s Met s	2208
ttc Phe	cte Let	g cc	g ca o Gl: 74	n va.	g gtg l Val	aag Lys	g tog s Sei	g gcg Ala 745	ı Arg	gtg y Val	g ato Met	aag Lys	750	n Ala	g gtg a Val	2256
gcg Ala	Cac His	75	1 110	t cce	c tto D Phe	ato E Ile	gag Glu 760	ı Glu	gaa Glu	aag Lys	g cgc	cag Gln 765	Ile	gcg Rala	g gcc Ala	2304
gcg Ala	999 Gly 770	GT	gat Asi	t gtg Val	g cgc l Arg	gco J Ala 775	a Lys	ggc Gly	aag Lys	ato Ile	gtg Val	Ile	gcc Ala	aco Thr	gtc Val	2352
aag Lys 785	Gly	gac Asp	gtg Val	g cac L His	: gac Э Авр 790	Ile	ggc Gly	aag Lys	aac Asn	atc Ile 795	Val	, tcg . Ser	gtg Val	gto Val	ttg Leu 800	2400
cag Gln	tgc Cys	aat Asn	aac Asr	Phe 805	: Glu	gtc Val	gtg Val	aac Asn	atg Met 810	Gly	gtg Val	atg Met	gtg Val	Pro 815	tgc Сув	2448
ATA	GIR	. 11e	820	GIn	. Гув	Ala	Lys	Авр 825	Glu	Asn	Ala	Asp	Met 830	Ile	ggc	2496
Leu	ser	835	Leu	lle	acg Thr	Pro	Ser 840	Leu	Glu	Glu	Met	Ala 845	Tyr	Val	Ala	2544
ser	850	met	Gin	Arg	gac Asp	Pro 855	Tyr	Phe	Arg	Glu	Arg 860	Ala	Met	Pro	Leu	2592
atg Met 865	ata Ile	ggc Gly	Gly 393	gcg Ala	acc Thr 870	acc Thr	agc Ser	cgg Arg	gtc Val	cat His 875	acg Thr	gcg Ala	gtc Val	aag Lys	atc Ile 880	2640
gcg Ala	ccc Pro	aac Asn	tac Tyr	gac Asp 885	ggt Gly	ccg Pro	gtg Val	atc Ile	tac Tyr 890	gtg Val	ccc Pro	gat Asp	gcc Ala	agc Ser 895	cgt Arg	2688
t <i>c</i> g Ser	gtc Val	ggc Gly	gtg Val 900	gcg Ala	acc Thr	agc Ser	ctc Leu	atg Met 905	tcc Ser	gac Asp	cag Gln	gcc Ala	ccg Pro 910	gcc Ala	tat Tyr	2736
ttg Leu	gcg Ala	gag Glu 915	ctg Leu	gcg Ala	cag Gln	gag Glu	tac Tyr 920	gag Glu	gat Asp	gtg Val	cgc Arg	cgc Arg 925	tgc Cys	cat His	gcc Ala	2784
1011	cgc Arg 930	aag Lys	gcg Ala	gtg Val	ccg Pro	ctg Leu 935	gtg Val	tcg Ser	ctg Leu	Ala	gag Glu 940	gcg Ala	cgc Arg	gcg Ala	gcg Ala	2832
rgc (Arg) 945	ccg Pro	cag Gln	atc Ile	gac Asp	tgg Trp 950	tcc Ser	ggc Gly	tac Tyr	Gln	ccg Pro 955	ccg Pro	cgc Arg	ccc Pro	aag Lys	ttc Phe 960	2880

	Lev	gg Gly	c cgg	g cg	g Ala 96!	a Phe	aag Lys	agc Ser	tac Tyr	gad Asj 970) Leu	gco Ala	gag Glu	ato Ile	gcg Ala 975	g cgc Arg	2928
	tat Tyr	ato	gad Asi	tgg Trj 980	p Gly	y ecg	ttc Phe	ttc Phe	Gln 985	Thi	tgg Trp	ago Ser	ctg Leu	Phe 990	Gly	ccg Pro	2976
	ttc Phe	Pro	gco Ala 995	a Ile	c ctg e Lev	gac Asp	Asp	aag Lys 1000	Val	, gtg Val	g ggc Gly	gag Glu	cag Gln 1005	gcg Ala	cgc	aag Lys	3024
	val	tac Tyr 1010	GIV	gaa Glu	a ggo a Gly	' Gla	gcc Ala 1015	atg Met	ctc	aag Lys	Arg	atc Ile 1020	Ile	gac Asp	999 Gly	cgc	3072
	tgg Trp 102	Leu	acc Thr	gcc Ala	ago Ser	ggc Gly 103	Val	gtc Val	ggc	ttc Phe	tat Tyr 103	Pro	gcc	aac Asn	cgc	gtc Val 1040	3120
	aat Asn	gac Asp	gaa Glu	gac Asp	atc Ile 1045	gag Glu	gtc Val	tac Tyr	Ala	gac Asp 1050	Glu	acg Thr	cgc Arg	Ser	gag Glu 1055	atg Met	3168
	ctg Leu	ttc Phe	Thr	tac Tyr 1060	Arg	aac Asn	ctg Leu	Arg	cag Gln 1065	cag Gln	ggc	gtc Val	Lys	cgc Arg 1070	gaa Glu	ggc Gly	3216
	gtc Val	ser	aac Asn 1075	aag Lys	tgc Cys	ctg Leu	Ala	gac Asp 1080	tac Tyr	atc Ile	gcg Ala	Pro	cgc Arg 1085	gac Asp	agc Ser	ggc Gly	3264
	ren	ctc Leu 1090	gac Asp	tac Tyr	atc Ile	ggc Gly	atg Met 1095	ttc Phe	gcc Ala	gtg Val	Thr	gcg Ala L100	ggc Gly	ctg Leu	ggc Gly	atc Ile	3312
	gag Glu 1105	гåв	aaa Lys	gag Glu	gcc Ala	gag Glu 1110	Phe	cag Gln	gcg Ala	gcg Ala	ctg Leu 1115	Asp	gac Asp	tac Tyr	tcc Ser	agc Ser 1120	3360
	atc Ile	atg Met	ctg Leu	Lys	tcg Ser 1125	ctg Leu	gcc Ala	gac Asp	Arg	ctg Leu 130	gcc Ala	gag Glu	gcg Ala	Phe	gcc Ala 135	gaa Glu	3408
	tgc Cys	atg Met	UIR	gcg Ala 1140	cgc Arg	gtg Val	cgc Arg	Arg	gac Asp 145	ctg Leu	tgg Trp	ggc Gly	Tyr	gcg Ala 150	gcg Ala	gac Asp	3456
	gag Glu	AIA	ctg Leu 155	tcc Ser	aac Asn	gat Asp	Glu	ctg Leu 160	atc Ile	gcc Ala	gag Glu	Lys	tac Tyr 165	agc Ser	ggc Gly	atc Ile	3504
	Arg	ccg Pro 170	gcg Ala	ccc Pro	ggc	tat Tyr 1	ccg Pro 175	gcc Ala	tgc Cys	ccg Pro	Glu	cac His 180	gtg Val	gtc Val	aag Lys	acg Thr	3552
4	gac Asp 1185	Leu	ttc Phe	cgc Arg	vai	ctg Leu 1190	gac (gcc (Ala /	gcc Ala	gac gac	gtc Val 1195	gga Gly	atg (Met (gag Glu	ctg Leu	acc Thr 1200	3600
•	gac	agc	tac	gcc	atg	ttc	ccg d	gcc (tcc .	aqc	atc	tca	aga 1	ttc	tat	ttc	3648

WO 03/087386 PCT/EP03/04010

139

3774

Asp Ser Tyr Ala Met Phe Pro Ala Ser Ser Val Ser Gly Phe Tyr Phe 1210

age cae ece gag teg cag tat tte aac gtg gge aac ate gge gee gae 3696 Ser His Pro Glu Ser Gln Tyr Phe Asn Val Gly Asn Ile Gly Ala Asp 1220 1225

cag ctg gcc gac tac gtg gcg cgc agc ggc cgc gcc gaa gag gac gtg 3744 Gln Leu Ala Asp Tyr Val Ala Arg Ser Gly Arg Ala Glu Glu Asp Val 1235 1240

cgc cgc acc ctg gcg ccg aac ctg ggc tag Arg Arg Thr Leu Ala Pro Asn Leu Gly 1250

<210> 32

<211> 1257

<212> PRT

<213> Bordetella pertussis

<220>

<221> unsure

<222> 69 .. 69

<223> All occurrences of Xaa indicate any amino acid

<220>

<221> unsure

<222> 93 .. 93

<223> All occurrences of Xaa indicate any amino acid

Val Pro Tyr Pro Arg Ile Pro Phe Pro Leu Ser Ala Tyr Thr His Gly

Gly Glu Phe Val Arg Gln Leu Asp Lys Arg Ile Leu Ile Leu Asp Gly 25

Ala Met Gly Thr Met Ile Gln Arg Tyr Lys Leu Gly Glu Ala Asp Phe

Arg Gly Glu Arg Phe Ala Glu His His Lys Asp Leu Lys Gly Asp Asn

Glu Leu Leu Ser Xaa Val Arg Pro Asp Val Ile Ala Glu Ile His Arg

Gln Tyr Leu Glu Ala Gly Ala Asp Val Ile Glu Thr Xaa Thr Phe Gly

Ala Thr Ser Ile Ala Gln Gly Asp Tyr Asp Leu Pro Glu Leu Ala Tyr 105

Glu Met Asn Leu Glu Ser Ala Arg Leu Ala Arg Ala Ala Cys Asp Ala

Tyr Ser Thr Pro Glu His Pro Arg Phe Val Ala Gly Ala Leu Gly Pro 135

Gln Pro Lys Thr Ala Ser Ile Ser Pro Asp Val Asn Asp Pro Gly Ala 155

WO 03/087386 PCT/EP03/04010

140

Arg Asn Val Thr Phe Asp Glu Leu Arg Ala Ala Tyr Val Glu Gln Leu 165 170 175

Asn Gly Leu Leu Asp Gly Gly Ile Asp Ile Val Leu Ile Glu Thr Ile

Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Val Glu Glu Ala
195 200 205

Phe Glu Ala Arg Gly Val Arg Leu Pro Val Met Ile Ser Gly Thr Val 210 220

Thr Asp Ala Ser Gly Arg Ile Leu Ser Gly Gln Thr Val Glu Ala Phe 225 230 235 240

Trp Asn Ser Val Arg His Ala Arg Pro Val Thr Ile Gly Leu Asn Cys 245 250 255

Ala Leu Gly Ala Ala Leu Met Arg Pro Tyr Val Ala Glu Leu Ser Lys 260 265 270

Ile Cys Asp Thr Tyr Val Cys Val Tyr Pro Asn Ala Gly Leu Pro Asn 275 280 285

Pro Met Ala Glu Thr Gly Phe Asp Glu Thr Pro Ala Asp Thr Ser Ala 290 295 300

Leu Leu Glu Glu Phe Ala Gln Ala Gly Leu Val Asn Met Ala Gly Gly 305 310 310 315 315

Cys Cys Gly Thr Thr Pro Glu His Ile Arg Ala Ile Ala Gly Lys Val

Ala Ala Leu Thr Pro Arg Ala Val Pro Glu Val Pro Val Lys Thr Arg

Leu Ser Gly Leu Glu Ala Leu Asn Ile Asp Asp Glu Thr Leu Phe Val 355 360 365

Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Lys Met Phe Ala Arg 370 380

Leu Val Arg Glu Glu Lys Tyr Asp Glu Ala Leu Ala Val Ala Arg Gln 385 390 395 400

Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Val Asn Met Asp Glu Ala

Met Leu Asp Ser Val Ala Cys Met His Arg Phe Leu Asn Leu Ile Ala
420 425 430

Ser Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys
435 440 445

Trp Glu Val Ile Glu Thr Gly Leu Lys Cys Val Gln Gly Lys Ala Val 450 460

Val Asn Ser Ile Ser Met Lys Glu Gly Glu Glu Pro Phe Arg His His 465 470 475 480

Ala Arg Leu Cys Arg Arg Tyr Gly Ala Ala Met Val Val Met Ala Phe
485 490 495

- Asp Glu Gln Gly Gln Ala Asp Ser Leu Glu Arg Arg Lys Glu Ile Cys 500 505 510
- Gly Arg Ala Tyr Arg Ile Leu Val Glu Glu Glu Gly Phe Pro Pro Glu 515 520 525
- Asp Ile Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Asp 530 535 540
- Glu His Asn His Tyr Ala Val Asp Phe Ile Glu Gly Ala Arg Trp Ile 545 550 555 560
- Arg Ala Asn Leu Pro His Ala Arg Ile Ser Gly Gly Ile Ser Asn Val 565 570 575
- Ser Phe Ser Phe Arg Gly Asn Glu Pro Met Arg Glu Ala Ile His Thr 580 585 590
- Val Phe Leu Tyr Tyr Ala Ile Glu Ala Gly Leu Thr Met Gly Ile Val 595 600 605
- Asn Ala Gly Gln Leu Gly Val Tyr Ala Asp Leu Ala Pro His Leu Arg 610 620
- Asp Leu Val Glu Asp Val Ile Leu Asp Arg Pro Glu Pro Val Gly Arg 625 630 635 640
- Ser Asp Ser Ala Asp Glu Arg Ser Pro Thr Glu Arg Leu Val Gln Phe 645 650 655
- Ala Glu Thr Val Lys Gly Ser Gly Ala Lys Lys Glu Glu Asp Leu Thr
 660 665 670
- Trp Arg Thr Gly Ser Val Glu Gln Arg Leu Ala His Ala Leu Val His
 675 680 685
- Gly Ile Thr Thr Phe Ile Val Glu Asp Thr Glu Glu Val Arg Gln Gln 690 695 700
- Val Ala Ala Arg Gly Gly Arg Thr Ile Glu Val Ile Glu Gly Pro Leu 705 710 715 720
- Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met 725 730 735
- Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val 740 745 750
- Ala His Leu Ile Pro Phe Ile Glu Glu Glu Lys Arg Gln Ile Ala Ala 755 760 765
- Ala Gly Gly Asp Val Arg Ala Lys Gly Lys Ile Val Ile Ala Thr Val 770 780
- Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser Val Val Leu 785 790 795 800
- Gln Cys Asn Asn Phe Glu Val Val Asn Met Gly Val Met Val Pro Cys 805 810 815
- Ala Gln Ile Leu Gln Lys Ala Lys Asp Glu Asn Ala Asp Met Ile Gly

825

- Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala Tyr Val Ala 835 840 845
- Ser Glu Met Gln Arg Asp Pro Tyr Phe Arg Glu Arg Ala Met Pro Leu 850 855 860
- Met Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val Lys Ile 865 870 875 880
- Ala Pro Asn Tyr Asp Gly Pro Val Ile Tyr Val Pro Asp Ala Ser Arg 885 890 895
- Ser Val Gly Val Ala Thr Ser Leu Met Ser Asp Gln Ala Pro Ala Tyr 900 905 910
- Leu Ala Glu Leu Ala Gln Glu Tyr Glu Asp Val Arg Arg Cys His Ala 915 920 925
- Asn Arg Lys Ala Val Pro Leu Val Ser Leu Ala Glu Ala Arg Ala Ala 930 935 940
- Arg Pro Gln Ile Asp Trp Ser Gly Tyr Gln Pro Pro Arg Pro Lys Phe 945 950 955 960
- Leu Gly Arg Arg Ala Phe Lys Ser Tyr Asp Leu Ala Glu Ile Ala Arg 965 970 975
- Tyr Ile Asp Trp Gly Pro Phe Phe Gln Thr Trp Ser Leu Phe Gly Pro 980 985 990
- Phe Pro Ala Ile Leu Asp Asp Lys Val Val Gly Glu Gln Ala Arg Lys 995 1000 1005
- Val Tyr Glu Glu Gly Gln Ala Met Leu Lys Arg Ile Ile Asp Gly Arg 1010 1015 1020
- Trp Leu Thr Ala Ser Gly Val Val Gly Phe Tyr Pro Ala Asn Arg Val 1025 1030 1035 1040
- Asn Asp Glu Asp Ile Glu Val Tyr Ala Asp Glu Thr Arg Ser Glu Met 1045 1050 1055
- Leu Phe Thr Tyr Arg Asn Leu Arg Gln Gln Gly Val Lys Arg Glu Gly 1060 1065 1070
- Val Ser Asn Lys Cys Leu Ala Asp Tyr Ile Ala Pro Arg Asp Ser Gly
 1075 1080 1085
- Leu Leu Asp Tyr Ile Gly Met Phe Ala Val Thr Ala Gly Leu Gly Ile 1090 1095 1100
- Glu Lys Lys Glu Ala Glu Phe Gln Ala Ala Leu Asp Asp Tyr Ser Ser 1105 1110 1115 1120
- Ile Met Leu Lys Ser Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1125 1130 1135
- Cys Met His Ala Arg Val Arg Arg Asp Leu Trp Gly Tyr Ala Ala Asp 1140 1145 1150

4	12	
ı	4.5	

Glu Ala Leu Ser Asn Asp Glu Leu Ile Ala Glu Lys Tyr Ser Gly Ile 1155 1160 1165

Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Val Val Lys Thr 1170 1175 1180

Asp Leu Phe Arg Val Leu Asp Ala Ala Asp Val Gly Met Glu Leu Thr 1185 1190 1195 1200

Asp Ser Tyr Ala Met Phe Pro Ala Ser Ser Val Ser Gly Phe Tyr Phe 1205 1210 1215

Ser His Pro Glu Ser Gln Tyr Phe Asn Val Gly Asn Ile Gly Ala Asp 1220 1225 1230

Gln Leu Ala Asp Tyr Val Ala Arg Ser Gly Arg Ala Glu Glu Asp Val 1235 1240 1245

Arg Arg Thr Leu Ala Pro Asn Leu Gly 1250 1255

<210> 33

<211> 3645

<212> DNA

<213> Chlorobium tepidum

<220>

<221> CDS

<222> (1) .. (3642)

<223> RCL00420

<400> 33

- gtg ctc gac ggg gcc atg ggc acc atg atc cag agg cat ggc ctc gac 48
 Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp
 1 5 10
- gaa cag gac tac cgg ggc gag cgt ttc gct tcg cat gac cat ccg ctg
 Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu
 20
 25
 30
- aag ggc aac aac gac ctt ctt gtc atc acc cgg ccc gac atc atc cgt
 Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg
 35 40 45
- tcg atc cac tgc gac ttc ctc gac gcg ggt gcg gac atc atc gag acc
 Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr
 50 55 60
- tgc acc ttc aac gcc aac ccg atc tcg cag tcg gac tac cag ttg cag

 Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln

 65 70 75 80
- gac ttg acc cgc gag ctg aac gtg gcg gcg gca aag ata gcc cgc tcg 288
 Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser
 85 90 95
- gca gcg gac gag ttc acc gca aag act ccc gac aag ccg cgt ttc gtg
 Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val
 100 105 110
- gcc ggt tcc atc gga ccg acc aac aag acg ctc tcg ctc tcg ccg gac 384

Al	a Gl	y Se 11	er Il 15	.e G1	y Pr	o Th	r As 12	n Ly O	s Th	r Le	u Se	r Lei 12		r Pro	qaA o	
V L	13	0	ou Pi	0 G1	уРп	13	5 5	a Va	l Th	r Pho	e Gl: 140	n Glu D	ı Me	t Va	c gat l Asp	432
14	11 1y	r in	r al	a GI	n Le	u Gli O	ı Gly	y Le	u His	3 Glv 159	ı Gly 5	/ Gly	v Va:	l Asp	t ctc Leu 160	480
Tel	ı ne	u va	T GI	u Th 16	r va. 5	l Phe	e Asi	Thi	170	ı Ası	з Сув	Lys	Ala	175		528
131	AI	a 11	18:	0 n G1	u Ty	r Ala	ı Val	185	Thr	Gly	Trp	Gln	Va]	Pro	gtg Val	576
, ME (. va.	19	5	y Tni	r val	l Val	200	Ala	Ser	Gly	Arg	Thr 205	Leu	Ser	ggc Gly	624
GII	. 210	Tn:	c Glí	1 Ala	a Phe	215	Ile	Ser	· Ile	Ser	His 220	Met	Pro	Ser	ctg Leu	672
225	sei	va.	ı Gış	Let	230	Cys	Ala	Leu	Gly	Ser 235	Lys	Gln	Met	Arg	Pro 240	720
Phe	116	GIU	ı Ala	245		Asn	Ile	Ala	Glu 250	Ser	Tyr	Val	Ser	Val 255	Tyr	768
PIO	ABII	MIA	260	reu	ccg Pro	Asn	Glu	Phe 265	Gly	Glu	Tyr	Asp	Asp 270	Ser	Pro	816
GIU	ıyı	275	Ala	Ala	cag Gln	He	Ala 280	Gly	Phe	Ala	Glu	Ser 285	Gly	Phe	Val	864
Wall	290	val	GIĀ	GIY	tgc Cys	Сув 295	Gly	Thr	Thr	Pro	Thr 300	His	Ile	Arg	Ala	912
305	,	Giu	WIG	vai	aag Lys 310	Tnr	Leu	Pro	Pro	Arg 315	Lys	Arg	Pro	Ala	Asn 320	960
гув	ulb	vai	ren	325	ctc Leu	Ser	Gly	Leu	Glu 330	Pro	Leu	Val	Val	Авр 335	Glu	1008
****	1111	Gly	340	116	aac Asn	Val	GIÀ	Glu 345	Arg	Thr	Asn	Val	Thr 350	Gly	Ser	1056
cgc Arg	aag Lys	ttc Phe 355	gcc Ala	cgc Arg	ctc Leu	11e	aag Lys 360	gag Glu	gcc Ala	aat Asn	Tyr	gac Asp 365	gaa Glu	gcg Ala	ctc Leu	1104

00.	37	0	a A	.y 61	.11 61	n va 37:	1 GI1 5	u Ası	a Gly	y Ala	38	n Vai 0	l Il	e As	c gtg p Val	1152
385	5	~ A.	p Gi	.u	3 9	0	n wal	sei	. G1/	1 Lys 395	s Vai	1 116	e Va	l Gl	ttc Phe 400	1200
500		u De	u 11	40	a se. 5	r GI	1 PTC	GIU	410	: Ala	Lys	s Val	l Pro	Val 415		1248
	. nol	, 56	42	0	e ii	e ser	. vai	425	Glu	Asn	Gl)	/ Lev	430	Cys	acc Thr	1296
-		43	5	. 11	c va.	L ASI	440	11e	ser	Leu	Lys	445	Gly	Glu	gag Glu	1344
Deu	450) F Dyl	9 61	n Wi	j Ala	455	Lys	Ile	Met	Gln	Tyr 460	Gly	Ala	Ala	gcg	1392
465	741	net	. MI	a PHE	470	GIU	GIN	GIA	Gln	Ala 475	Asp	Ser	Leu	His	480	1440
Æg	116	GIC	. 116	485	ser	cgc Arg	Ala	Tyr	Lys 490	Ile	Leu	Thr	Glu	Glu 495	Val	1488
Uly	rne	PIU	500	GIU	Asp	atc Ile	Ile	Phe 505	qaA	Pro	Asn	Val	Leu 510	Thr	Val	1536
7124		515	116	чар	GIU	cac His	Asn 520	Asn	Tyr	Ala	Leu	Asp 525	Phe	Ile	Glu	1584
	530	~-3	11.0	116	гуя	cag Gln 535	Asn	Leu	Pro	His	Ala 540	Lys	Val	Ser	Gly	1632
545		501	VOII	val	550	ttc Phe	ser	Phe	Arg	G1y 555	Asn	Glu	Pro	Val	Arg 560	1680
				565	ALG	ttc Phe	reu	1yr	H18 570	Ala	Ile	His	Ala	Gly 575	Leu	1728
gac (nie C	GIY	580	Val	ASI	Ala .	Ala (G1n 1 585	Leu (Gly	Ile	Tyr	Glu 590	Glu	Ile	1776
gac (595	Deu	Deu	val	lyr	600	ilu /	gap ,	Val 1	Leu	Leu . 605	Asn	Arg	Arg	1824
gac	gac	gcc	acc	gag	cgg	ctc	gtg g	gcg t	tc g	gct g	gaa .	acg .	atc	cgc	gac	1872

										146				•		
ВA	P As 61	р А. 0	la Tì	ar Gl	u Ar	g Lei 61	u Va 5	l Ala	a Pho	e Al	a Gl		r Il	e Arg	qaA g	
99 G1 62	A. GI	c ga y G]	a aa lu Ly	ng go /8 Al	c gag a Gli 630	u Ala	a Ly	g aad 9 Ası	e gee	gaa Glu 63!	ı Tr	p Arg	aac Ası	gco n Ala	ccg Pro 640	1920
va.	I GI	u GI	u Ar	g Le 64	u Ly: 5	s His	s Ala	a Leu	2 Val	l Lys	3 Gly	/ Ile	val	l Asp 655		1968
ate Ile	e As	c ga p Gl	g ga u As 66	p Th	c gaa r Glu	a gag ı Glu	g gco L Ala	c cgc Arg 665	, Glr	cto Lev	tac Tyr	e cce	g agt Ser 670	Pro	ctg Leu	2016
GI	ı va.	67	e G1 5	u G1;	y Pro	Leu	680	: Asn	Gly	Met	. Asr	His 685	Va]	Gly	gac Asp	2064
rer	690) ; AI	a GI	u GI	c aag y Lys	Met 695	Phe	. Leu	Pro	Gln	700	. Val	Lys	Ser	Ala	2112
705	, vai	. ме	с гу	s Arg	tcg Ser 710	Val	Ala	Ala	Leu	715	Pro	Tyr	Ile	Glu	Glu 720	2160
GIU	Lys	s Se	c r.y	725		Asp	Thr	Ser	Ala 730	Lys	Ala	Lys	Val	Leu 735	Leu	2208
ATA	ing	va.	740) G GIA	gac Asp	Val	His	Asp 745	Ile	Gly	Lys	Asn	Ile 750	Val	Ser	2256
vai	vai	755	;	ι Сув	aac Asn	Asn	Phe 760	qeA	Val	Ile	Asp	Ile 765	Gly	Val	Met	2304
MEC	770	Сув	ASP	, гав	att Ile	775	Glu	Ala	Leu	Ala	Glu 780	His	Lys	Pro	yeb	2352
gtg Val 785	ctc Leu	ggo	cto Leu	tcc Ser	ggc Gly 790	ctc Leu	atc Ile	acc Thr	ccg Pro	tcg Ser 795	ctc Leu	gaa Glu	gag Glu	atg Met	gcg Ala 800	2400
cac His	gtg Val	gcc	aaa Lys	gag Glu 805	atg Met	gag Glu	cgg Arg	ctc Leu	ggc Gly 810	atg Met	aac Asn	att Ile	ccg Pro	ctc Leu 815	atc Ile	2448
atc Ile	ggc	ggc	gcg Ala 820	acc Thr	acc Thr	tcg Ser	aag Lys	gtg Val 825	cac His	acg Thr	gcg Ala	gtg Val	aaa Lys 830	ctc Leu	gcg Ala	2496
ccc Pro	tgc Cys	tac Tyr 835	ccc Pro	agc Ser	Gly	Ala	gta Val 840	gta Val	cac His	gtg Val	ctc Leu	gac Asp 845	gcc Ala	tcg Ser	cgc Arg	2544
agc Ser	gtg Val 850	ccg Pro	gtg Val	gtc Val	agc Ser	aac Asn 855	ctc Leu	tgc Cys	aac Asn	Pro	gcc Ala 860	cag Gln	cgc Arg	gac Asp	agc Ser	2592

tat Tyr 865	Ile	gcg Ala	gcg Ala	ctg Leu	aag Lys 870	gat Asp	gag Glu	cag Gln	gag Glu	gcg Ala 875	Met	cgc Arg	aag Lys	agc Ser	cac His 880	2640
gcc	gag	egc Arg	atg Met	gcg Ala 885	gca Ala	aaa Lys	aag Lys	tac Tyr	gtc Val 890	Ser	ctc Leu	Aap	gcc Ala	gcc Ala 895	Arg	2688
gac Asp	Asn	cgc Arg	ctc Leu 900	acc Thr	att	gac Asp	tgg Trp	gag Glu 905	gcc Ala	gaa Glu	acc Thr	atc Ile	gac Asp 910	aag Lys	ccc Pro	2736
gcc Ala	cag Gln	act Thr 915	ggc Gly	gtc Val	acc Thr	gtg Val	ctg Leu 920	gag Glu	gat Asp	gtc Val	acc Thr	gtc Val 925	Gly	gcg Ala	ctc Leu	2784
Arg	Pro 930	Tyr	Ile	Asp	tgg Trp	Ala 935	Xaa	Phe	Phe	Trp	Ser 940	Trp	Glu	Leu	His	2832
Gly 945	Val	Tyr	Pro	Gln	att Ile 950	Leu	Glu	Asp	Glu	Lys 955	Val	Gly	Glu	Glu	Ala 960	2880
Thr	Lys	Leu	Phe	Asn 965	gac Asp	Ala	Thr	Ala	Leu 970	Leu	Asp	Arg	Ile	Asp 975	Ser	2928
Glu	Lys	Leu	Leu 980	Gly	atc Ile	Lys	Gly	Val 985	Ala	Gly	Ile	Phe	Pro 990	Ala	Asn	2976
Ser	Ile	Gly 995	qaA	qaA	atc Ile	Phe]	Val 1000	Tyr	Ala	Двр	Asp I	Glu 1005	Arg	Ser	Ile	3024
lle 1	Arg 1010	Thr	Val	Leu		Thr 015	Leu	Arg	Gln	Gln 1	Gly 1020	Glu	Lys	His	Gly	3072
G1u 1025	Ala	Asn	Leu	Ala	ctg Leu 1030	Ala	Asp	Phe	Val	Ala 1035	Pro	Arg	Glu	Ser	Gly 1040	3120
Val	Asn	Asp	Trp 1	11e 045	ggc Gly	Сув	Phe	Thr 1	Val 050	Thr	Ala	Gly	Leu 1	Gly .055	Ile	3168
Gln	Asn	Leu 1	Leu . 060	Asp	gag Glu	Phe	Thr 1	Ala 065	Glu	Asn	Asp	Asp 1	Tyr 070	His	Arg	3216
Ile	Met 1	Thr •	Gln /	Ala	ctc (Leu)	Ala 1	Asp 2	Arg	Leu	Ala	Glu 1	Ala 085	Phe	Ala	Glu	3264
Met 1	Leu 090	His (Glu 1	Lys		Arg . 095	Arg (3lu	Leu	Trp 1	Gly 100	Tyr	Ala	Pro	Gly	3312
gaa	atc	ctc	ggc a	aac	gaa q	gag	ctg a	atc (gcc	gaa	aag	tac	cga	ggc	atc	3360

148	
Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile 1105 1110 1115 1120	•
cgc ccc gcc ccc ggc tac ccc gcc tgc ccg gat cac acc gaa aag gca Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala 1125 1130 1135	3408
atc atc ttc gac ctg ctc aac gct gaa gcg gcc acc ggc gtc acg ctg Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu 1140 1145 1150	3456
acg gaa act ttc gcg atg aac ccc gca gcc tca gtc tgc ggc ctc tac Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr 1155 1160 1165	3504
ttc gcc aac ccg gcc tcg aaa tac ttc gta ctc ggc aag att ggt aag Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys 1170 1175 1180	3552
gat cag gtc gaa gac tac gcc aac cgc aaa ggg ctg gaa gta gca gaa Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu 1185 1190 1195 1200	3600
gcc gag aag tgg ctc gcg ccc tcg ctg aac tac gat cca gcg Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala 1205 1210	3642
taa	3645
<210> 34 <211> 1214 <212> PRT <213> Chlorobium tepidum	
<220> <221> unsure <222> 936 936 <222> All occurrences of Xaa indicate any amino acid	
<400> 34	
Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp 1 5 10 15	
Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu 20 25 30	
Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg 35 40 45	
Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr 50 55 60	
Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln 65 70 75 80	
Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser 85 90 95	
Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val	

WO 03/087386 PCT/EP03/04010

149

Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp 115 120 125

Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp. 130 135 140

Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu 145 155 160

Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu 165 170 175

Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val 180 185 190

Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly 195 200 205

Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu 210 220

Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro 225 230 235 240

Phe Ile Glu Ala Leu Ser Asn Ile Ala Glu Ser Tyr Val Ser Val Tyr 245 250 255

Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp Asp Ser Pro 260 265 270

Glu Tyr Met Ala Ala Gln Ile Ala Gly Phe Ala Glu Ser Gly Phe Val 275 280 285

Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Thr His Ile Arg Ala
290 295 300

Ile Ala Glu Ala Val Lys Thr Leu Pro Pro Arg Lys Arg Pro Ala Asn 305 310 315 320

Lys His Val Leu Arg Leu Ser Gly Leu Glu Pro Leu Val Val Asp Glu 325 330 335

Thr Thr Gly Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser 340 345 350

Arg Lys Phe Ala Arg Leu Ile Lys Glu Ala Asn Tyr Asp Glu Ala Leu 355 360 365

Ser Ile Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile Asp Val 370 380

Asn Leu Asp Glu Gly Met Leu Asp Ser Glu Lys Val Ile Val Glu Phe 385 390 395 400

Leu Asn Leu Ile Ala Ser Glu Pro Glu Ile Ala Lys Val Pro Val Met
405 410 415

Ile Asp Ser Ser Lys Trp Ser Val Ile Glu Asn Gly Leu Arg Cys Thr 420 425 430

Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu 435 440 445

- Leu Phe Lys Glu Arg Ala Arg Lys Ile Met Gln Tyr Gly Ala Ala Ala Ala 450

 Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Ser Leu His Arg 480
- Arg Ile Glu Ile Cys Ser Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val 485 490 495
- Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val 500 505 510
- Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile Glu 515 520 525
- Ser Val Arg Trp Ile Lys Gln Asn Leu Pro His Ala Lys Val Ser Gly 530 540
- Gly Ile Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg 545 550 555 560
- Glu Ala Met His Thr Ala Phe Leu Tyr His Ala Ile His Ala Gly Leu 565 570 575
- Asp Met Gly Ile Val Asn Ala Ala Gln Leu Gly Ile Tyr Glu Glu Ile 580 585 590
- Asp Pro Glu Leu Leu Val Tyr Val Glu Asp Val Leu Leu Asn Arg Arg 595 600 605
- Asp Asp Ala Thr Glu Arg Leu Val Ala Phe Ala Glu Thr Ile Arg Asp 610 620
- Gly Gly Glu Lys Ala Glu Ala Lys Asn Ala Glu Trp Arg Asn Ala Pro 625 630 635 640
- Val Glu Glu Arg Leu Lys His Ala Leu Val Lys Gly Ile Val Asp Tyr 645 650 655
- Ile Asp Glu Asp Thr Glu Glu Ala Arg Gln Leu Tyr Pro Ser Pro Leu 660 665 670
- Glu Val Ile Glu Gly Pro Leu Met Asn Gly Met Asn His Val Gly Asp 675 680 685
- Leu Phe Ala Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala 690 695 700
- Arg Val Met Lys Arg Ser Val Ala Ala Leu Ile Pro Tyr Ile Glu Glu 705 710 715 720
- Glu Lys Ser Lys Asn Cys Asp Thr Ser Ala Lys Ala Lys Val Leu Leu 725 730 735
- Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser 740 745 750
- Val Val Leu Ala Cys Asn Asn Phe Asp Val Ile Asp Ile Gly Val Met 755 760 765
- Met Pro Cys Asp Lys Ile Leu Glu Ala Leu Ala Glu His Lys Pro Asp

770

775

- Val Leu Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala 785 790 795 800
- His Val Ala Lys Glu Met Glu Arg Leu Gly Met Asn Ile Pro Leu Ile 805 810 815
- Ile Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Leu Ala 820 825 830
- Pro Cys Tyr Pro Ser Gly Ala Val Val His Val Leu Asp Ala Ser Arg 835 840 845
- Ser Val Pro Val Val Ser Asn Leu Cys Asn Pro Ala Gln Arg Asp Ser 850 855 860
- Tyr Ile Ala Ala Leu Lys Asp Glu Gln Glu Ala Met Arg Lys Ser His 865 870 875 880
- Ala Glu Arg Met Ala Ala Lys Lys Tyr Val Ser Leu Asp Ala Ala Arg 885 890 895
- Asp Asn Arg Leu Thr Ile Asp Trp Glu Ala Glu Thr Ile Asp Lys Pro 900 905 910
- Ala Gln Thr Gly Val Thr Val Leu Glu Asp Val Thr Val Gly Ala Leu 915 920 925
- Arg Pro Tyr Ile Asp Trp Ala Xaa Phe Phe Trp Ser Trp Glu Leu His 930 935 940
- Gly Val Tyr Pro Gln Ile Leu Glu Asp Glu Lys Val Gly Glu Glu Ala 945 950 955 960
- Thr Lys Leu Phe Asn Asp Ala Thr Ala Leu Leu Asp Arg Ile Asp Ser 965 970 975
- Glu Lys Leu Gly Ile Lys Gly Val Ala Gly Ile Phe Pro Ala Asn 980 985 990
- Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser Ile 995 1000 1005
- Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His Gly 1010 1015 1020
- Glu Ala Asn Leu Ala Leu Ala Asp Phe Val Ala Pro Arg Glu Ser Gly 1025 1030 1035 1040
- Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Ile 1045 1050 1055
- Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Arg 1060 1065 1070
- Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1075 1080 1085
- Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly 1090 1095 1100

/EP03/04010

													•			
	W	O 03/	08738	36												PCT
										152		•		•		
Gl:	u Il OS	e Leı	ı Gl	y Ası	1 Glu	ı Glu	ı Le	u Ile	e Ala	152 a Glu 1115	ı Lyı	з Туг	r Arg	Gly	/ Ile 1120	
Ar	g _, Pro	o Ala	a Pro	Gly 1125	y Tyn	Pro	Ala	а Суя	9 Pro	о Авр	Hia	Thi	: Glu	Lya 1135	Ala	
Ile	e Ile	e Phe	: Авр 1140	Lev)	ı Lev	ı Asr	Ala	1 Glu	ı Ala	Àla	Thi	Gly	/ Va] 1150		Leu	
Thi	Glu	1 Thr 1155	Phe	: Ala	Met	Asn	Pro	Ala	. Ala	Ser	Val	. Суа 1165		Leu	Tyr	
Phe	Ala 1170	Asn	Pro	Ala	Ser	Lys 1175	Туг	Phe	val		Gly 1180		Ile	Gly	Lys	
Asp 118	Gln 5	Val	Glu	Asp	Tyr 1190	Ala	Asn	Arg	l Lys	Gly 1195	Leu	Glu	Val		Glu 1200	
Ala	Glu	Lys	Trp	Leu 1205	Ala	Pro	Ser		Asn 1210		Asp	Pro	Ala			
<21	0> 3	5					,									
	1> 3															
	2> D 3> D		cocc	us r	adio	dura	ns									
<22	0 >													3.		
<22	1> C															
		1) DR026		4)												
			,,,											٠,		
	0 > 3	-														
Met 1	Ser	His	His	Pro 5	gaa Glu	gcg Ala	ccg Ser	gct Ala	Ser 10	gcc	aat Asn	ccg Pro	tcc Ser	atc Ile 15	aac Asn	4
cat His	caa Gln	PIO	tcc Ser	acc Thr	atc Ile	acc Thr	gag Glu	gcc Ala	ġcc Ala	cgc Arg	cag Gln	cgc Arg	atc Ile	ctg Leu	att Ile	9(

ctc gac ggc gcc tgg ggt acg cag ctt cag cga gcc aac ctc acc gaa Leu Asp Gly Ala Trp Gly Thr Gln Leu Gln Arg Ala Asn Leu Thr Glu 40

geg gac tte ege tgg gac gaa gee gac eee acg egg atg tae egg gge Ala Asp Phe Arg Trp Asp Glu Ala Asp Pro Thr Arg Met Tyr Arg Gly

aac ttc gac ctg ctg caa ctg acc aag cct gac gtg att cgc gcc gtg

Asn Phe Asp Leu Leu Gln Leu Thr Lys Pro Asp Val Ile Arg Ala Val

cac ege gee tat the gag gee gga geg gae ate gee age ace aat ace

His Arg Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ala Ser Thr Asn Thr

ttc aac tcc acg acc atc tcg cag gcg gat tac ggc acc gag gca ctg Phe Asn Ser Thr Thr Ile Ser Gln Ala Asp Tyr Gly Thr Glu Ala Leu

gcc tac gcc atg aac cgc gag ggg gca agg ctg gcc cgc gaa gtc gcc

105

55

85

100

50

65

30

240

										153						
Ala	а Ту	r Ala 11	a Mei	t As:	n Ar	g Gl	120	/ Ala	Arg	J Leu	ı·Ala	Arg 125		Va]	Ala	
yei Yei	ga Gli	u Phe	e Glu	g gcg 1 Ala	g cgo	g gad g Asp 135	o Gly	aaa Lys	aag Lys	g cgc B Arg	tgg Trp 140	Val	gcg	G13	agt Ser	432
gto Val 145	Gl	Pro	c acc	aac Ası	c cgc Arg 150	Th:	gcg Ala	acc Thr	ctt Leu	tct Ser 155	Pro	gac Asp	gtg Val	gag Glu	cgg Arg 160	480
Pro	gag Glu	j tto 1 Phe	cgc Arg	aac Asr 165	ı Val	g acc	tac Tyr	Asp Asp	gac Asp 170	Leu	gtg Val	gcg Ala	gcg	tac Tyr 175		528
gag	gco Ala	ato Ile	t acc Thr 180	Gly	tt <u>c</u> Lev	, atg Met	gaa Glu	ggt Gly 185	Gly	gcg Ala	gac Asp	ctg Leu	ctg Leu 190	Leu	att Ile	576
gaa Glu	acg Thr	gtg Val 195	Phe	gac Asp	acg Thr	ctg Leu	aac Asn 200	Ala	aaa Lys	gcc Ala	gcg Ala	ctg Leu 205	ttt Phe	gcc Ala	gcg Ala	624
Gin	210	Val	Phe	Ala	gcg Ala	Gln 215	Gly	Arg	Glu	Leu	Pro 220	Val	Met	Leu	Ser	672
ggc Gly 225	Thr	atc Ile	acc Thr	Asp	gcc Ala 230	tcg Ser	ggc	cgc Arg	acg Thr	ctg Leu 235	agc Ser	Gly 999	cag Gln	acg Thr	ccc Pro 240	720
gaa Glu	gcc Ala	ttc Phe	gcg Ala	gtg Val 245	agc Ser	acc Thr	gag Glu	cac His	gcc Ala 250	ggc	ctc Leu	ttt Phe	tcg Ser	ctg Leu 255	ggc	768
ctg Leu	aac Asn	tgc Cys	gcg Ala 260	ctg Leu	ggc	gcc Ala	gac Asp	ctg Leu 265	ctg Leu	cgg Arg	ccc Pro	cac His	ctg Leu 270	cgc Arg	gca Ala	816
att Ile	gcg Ala	gcg Ala 275	aac Asn	acg Thr	gag Glu	gcg Ala	ctg Leu 280	gtg Val	tcg Ser	gtt Val	cac His	ccc Pro 285	aac Asn	gcg Ala	ggc Gly	864
ctc Leu	ccc Pro 290	aac Asn	gcc Ala	ttc Phe	G1 y 999	gaa Glu 295	tac Tyr	gac Asp	gaa Glu	acg Thr	ccc Pro 300	gaa Glu	cac His	acg Thr	gcg Ala	912
gcg Ala 305	gtg Val	ctg Leu	gcc Ala	gac Asp	ttc Phe 310	gcc Ala	cgc Arg	gag Glu	ggg ggg	ctg Leu 315	gtc Val	aac Asn	atc Ile	gtg Val	ggc Gly 320	960
ggc Gly	tgc Cys	tgc Cys	ggc Gly	acc Thr 325	aca Thr	ccc Pro	gag Glu	His	atc Ile 330	aaa Lys	gcg Ala	att Ile	gcg Ala	gag Glu 335	gcg Ala	1008
gtg Val	aag Lys	gac Asp	att Ile 340	ccc Pro	ccg Pro	cgc Arg	Gln	gcg Ala 345	ctg Leu	caa Gln	ctg Leu	ccg Pro	cct Pro 350	tac Tyr	ctg Leu	1056
cgc Arg	ctc Leu	agc Ser 355	ggc Gly	ctc Leu	gaa Glu	Ala	ttc Phe 360	acc Thr	ctg Leu	acg Thr	Pro	gaa Glu 365	acc Thr	aac Asn	ttc Phe	1104

val	370	n va D	I GI	A GT	ı Arç	375	. Asr	ı Va	l Th:	r Gly	7 Sei 380	r Pro	Ly	B Phe	e Ser	1152
aag Lys 385	N ATE	a Il	t cto	g gco 1 Ala	990 G Gly 390	/ Asp	tac Tyr	: gad	gca Ala	a ggg a Gly 395	/ Let	aag 1 Lys	l Ile	e Ala	cgc Arg 400	1200
cag Gln	Cag Glr	g gtg 1 Va:	g aco	aac Asr 405	ı Gly	gcg Ala	caa Gln	ato	yal Val	yst y	ato Ile	aac Asn	tto Phe	gac 2 Asp 415	gag Glu	1248
61À 888	atg Met	Cto Lev	gad Asp 420	Gly	gaa Glu	gga Gly	gcg Ala	atg Met 425	: Val	aag Lys	tto Phe	ctc Leu	aac Asr 430	Leu	ctc Leu	1296
gcc Ala	G1y	gag Glu 435	Pro) Asp	atc Ile	tcg Ser	cgc Arg 440	gtg Val	Pro	ctg Leu	atg Met	Leu 445	Asp	tcg Ser	tcc Ser	1344
aag Lys	tgg Trp 450	GIU	l Ile	ctg Leu	gaa Glu	gcg Ala 455	G1y 999	ctg Leu	cgg	cgg Arg	gtg Val 460	Gln	ggo	aag Lys	gca Ala	1392
gtc Val 465	gtc Val	aac Asn	tcc Ser	atc Ile	tcg Ser 470	ctc Leu	aag Lys	gac Asp	ggc	gag Glu 475	Ala	agg Arg	ttt Phe	ctg Leu	gaa Glu 480	1440
Arg	AIA	Arg	Leu	Leu 485	Arg	cgc Arg	Tyr	Gly	Ala 490	Ala	Ala	Val	Val	Met 495	Ala	1488
ttc Phe	gac Asp	gaa Glu	cag Gln 500	gga Gly	cag Gln	gcc Ala	Asp gac	aac Asn 505	ctc Leu	gcc Ala	cga Arg	cgc Arg	cgg Arg 510	gag Glu	att Ile	1536
Leu	GIĄ	515	Ala	ıyr	Arg	ctg Leu	Leu 520	Thr	Glu	Gln	Ala	Asp 525	Phe	Pro	Pro	1584
cag Gln	gac Asp 530	atc Ile	att Ile	ttc Phe	gac Asp	ccc Pro 535	aac Asn	gtg Val	ctg Leu	acc Thr	gtt Val 540	gcc Ala	acc Thr	ggc Gly	atc Ile	1632
gag Glu 545	gaa Glu	cac His	gac Asp	cgc Arg	tac Tyr 550	gcg Ala	ctg Leu	gac Asp	ttt Phe	atc Ile 555	gag Glu	gcg Ala	acg Thr	cgc Arg	tgg Trp 560	1680
att Ile	aaa Lys	gaa Glu	aac Asn	ctg Leu 565	ccg Pro	gcg Ala	gcg Ala	aag Lys	gtg Val 570	tcg Ser	ggc Gly	999 999	att Ile	tcc Ser 575	aac Asn	1728
gtc Val	tcg Ser	ttc Phe	agc Ser 580	ttc Phe	cgg Arg	ggc Gly	Asn	aac Asn 585	cac His	gtg Val	cgc Arg	gag Glu	gcg Ala 590	atg Met	cac His	1776
gcg (val	ttt Phe 595	ctg Leu	tac Tyr	cac His	Ala :	atc Ile 600	cgc Arg	gcc Ala	999 Gly	Leu	gac Asp 605	atg Met	ggc Gly	atc Ile	1824
gtg a	aac	gcg	999	atg (ctg	gcg (gtg	tac	gag	gac	atc	gag	ccg	gag	ctg	1872

	17-	٠.								_		100						
	va	6 1	5n 10	Alč	1 G1	у ме	t Le	u Al. 61	a Va 5	1 ту	r Gl	u As	p Il 62		u Pr	o Gl	u Leu	
	62	9 G 5	ΙU	A16	ı va	1 G1	u As 63	p Va. O	1 11	e Le	u Ala	63	g Arg	g Pro	O Asj	p Ala	c acc a Thr 640	1920
	GI	u A.	ŗg	rec	те	64:	r Lei	n Ale	a As	p Ar	g Ty: 650	r Lyı	s Asj	o Ile	E Ly	659	=	1968
	361	Г А.	la	ATA	661	n se:	r Ala	a TrŢ	Arg	669 669	p Leu 5	ı Pro	o Val	l Glr	670	ı Arg	g ctg g Leu	2016
	vr.	, n.	. 5	675	TIEC	ı va.	l GII	ı Gıy	680)	a Asp) Hig	y Va]	l Asp 1883	Glu	ı Asp	gcc Ala	2064
	GIU	69	0	АТА	172	GII	ı Glu	695	Gly	Ser	Pro	Leu	700	Val	Ile	Glu	ggc	2112
	705	Le	u	met	мвр	GIY	710	Asn	Val	Val	. Gly	715	Leu	Phe	Gly	Ala	999 Gly 720	2160
	гур	ME	٠.	rne	ren	725	GIN	Val	Val	Lys	730	Ala	Arg	Val	Met	Lys 735		2208
	VIG	Va		HIA	740	Leu	Tnr	ccc Pro	Tyr	Leu 745	Glu	Ala	Glu	Lys	Ala 750	Glu	Ser	2256
	Ser	5e.	7	755	GIY	ràe	Val	ctg Leu	Leu 760	Ala	Thr	Val	Lys	Gly 765	Asp	Val	His	2304
	veb	770	ב כ ס	ыл	гув	ABD	TIE	gtg Val 775	Gly	Val	Val	Leu	Ala 780	Сув	Asn	Gly	Tyr	2352
	785	va.	. 1	nr	Авр	ren	790	gtg Val	Met	Val	Pro	Gly 795	Glu	Lys	Ile	Leu	да Д 008	2400
,	GIU	WIS	ıG	iu .	Arg	805	GIY	gcc Ala	Asp	Val	Ile 810	Gly	Leu	Ser	Gly	Leu 815	Ile	2448
i	acg Thr	Pro	t S	er .	tta Leu 820	gac Asp	gaa Glu	atg Met	Val	aac Asn 825	gtg Val	gcc Ala	cgc Arg	gag Glu	atg Met 830	acg Thr	cgc Arg	2496
2	cgg Arg	ggc Gly	•	tg a al 1 35	aaa Lys	act Thr	cca Pro	ctg Leu	ctg Leu 840	atc Ile	ggc Gly	ggc	gcg Ala	acg Thr 845	acc Thr	agc Ser	cgg A rg	2544
9	41 a	cac His 850	17	cg g hr 1	gcg Na	gtc Val	rae	att (lle) 855	gac Asp	ccg Pro	gcc Ala	Tyr	gac Asp 860	G1y 999	acg Thr	gta Val	gtg Val	2592

865	val	Leu	. Asp) Ala	870	Arg	, Ala	Val	. Thr	7 Val 875	Thr	Asn	Asp	Leu	880	· 2640
acc Thr	gac Asp	gag Glu	gcc	gcc Ala 885	Tyr	gct	Gly Gly	cgc Arg	ytg Val 890	Gln	ggc	gag Glu	tat Tyr	gac Asp 895	acc	2688
rea	Arg	GIU	900	H18	Gly	Glu	Arg	Gln 905	Val	Arg	Leu	Ile	Ala 910	Leu	gca Ala	2736
GIU	Ala	915	AJA	Arg	Ala	Pro	Gln 920	Leu	Ser	Ala	Ala	Val 925	Pro	Pro	gcg Ala	2784
PIO	930	Asp	Leu	GIÀ	Arg	Gln 935	Val	Val	Glu	Gln	Pro 940	Ile	Ala	Glu	ctg Leu	2832
945	PIO	Pne	116	Ąsp	950	Thr	Pro	Phe	Phe	atc Ile 955	Ala	Trp	Glu	Met	Lув 960	2880
GIÅ	116	Tyr	Pro	Gly 965	Ile	Leu	Thr	Хвр	Pro 970	ctg Leu	Arg	Gly	Glu	Glu 975	Ala	2928
ÀIG	гув	ren	980	Ala	Asp	Ala	Gln	Ala 985	Leu	ctg Leu	Glu	Gln	Val 990	Ile	Ala	2976
АВР	GIÀ	995	Leu	Arg	Ala	Arg 1	Gly .000	Val	Ile	61Å 888	Leu 1	Trp 1005	Pro	Ala	His	3024
1	010	wab.	116	val	Leu 1	015	qeA	Ala	Ala		Gly 1020	Arg	Gly	Glu	Thr	3072
1025	АБР	Pne	GIU	inr	н18 1030	Glu	Leu	Ala	Ala	1035	Arg	Glu	Pro	Leu	Pro 1040	3120
aac Asn I	MEC .	PIO A	arg 1	Leu 045	H18	Thr	Leu .	Arg 1	Gln 050	Gln	Arg	Авр	Gln 1	Thr 055	Thr	3168
Pro	ASA .	10	060	Leu .	Ala	Asp :	Phe 1	Val . 065	Ala	Glu	Gly	Gly .	Asp 070	His	Ile	3216
ggc g	10)75	41a :	rnr i	Ala :	11e 1	Phe (3ly /	Ala	Glu	Glu :	Leu 1 085	Ala	Gln	Gln	3264
	91u <i>f</i> 990	tia (in i	118	4sp /	18p 7	Tyr 1	lsn :	Ser	Ile 1	Leu ' 100	Val 1	Lys .	Ala	Val	3312
gcc g	jac C	ga c	rg g	gce (gag g	gcc t	tt g	lcc d	gag (aag	ctg	cac (egc (gac	gtg	3360

Ala Asp 1105	Arg	Leu	Ala	Glu 1110	Ala D	Phe	Ala		57 Lys 111!		His	Arg	Дар	Val 1120
cgc gtg Arg Val	cgg Arg	H18	tgg Trp 1125	ggt Gly	tac Tyr	gcc Ala	Glu	ggc Gly 1130	gag Glu	gcg Ala	ctc Leu	qaA	aac Asn l135	acc Thr

- gac ctc atc aag gag cgc tat cag ggc atc cgc cct gcg ccc ggc tac 3456 Asp Leu Ile Lys Glu Arg Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr 1140 1145 1150
- Pro Ala Gln Pro Asp His Thr Glu Lys Arg Thr Leu Phe Glu Leu Leu
 1155
 1160
 1165
- gac gcg gaa agc atc ggc ctg cgc ctc acc gag tcg tgt gcc atg acc 3552
 Asp Ala Glu Ser Ile Gly Leu Arg Leu Thr Glu Ser Cys Ala Met Thr
 1170 1175 1180
- ccg gcg gcg gcg gtg tcg ggg ctg tac ttc gcg cat ccg gag gcc cgt 3600 Pro Ala Ala Ala Val Ser Gly Leu Tyr Phe Ala His Pro Glu Ala Arg 1185 1190 1195 1200
- tat ttc gca gtg ggc cgc atc ggg cgc gac cag gtg gag aac tac gcc 3648 Tyr Phe Ala Val Gly Arg Ile Gly Arg Asp Gln Val Glu Asn Tyr Ala 1205 1210 1215
- gcc cgt aag ggt tgg act gtg cag gaa gcc gag cgc tgg ctg ggg ccg Ala Arg Lys Gly Trp Thr Val Gln Glu Ala Glu Arg Trp Leu Gly Pro
 1220 1225 1230
- ctg ctg gcg tac agc gcc ggg ccg ggg cca gaa gca agc cag aaa gcc 3744 Leu Leu Ala Tyr Ser Ala Gly Pro Gly Pro Glu Ala Ser Gln Lys Ala 1235 1240 1245
- ctc ggc gca gag ctg aca gga gcg caa tcg tga 3777 Leu Gly Ala Glu Leu Thr Gly Ala Gln Ser 1250 1255

<210> 36

<211> 1258

<212> PRT

<213> Deinococcus radiodurans

<400> 36

Met Ser His His Pro Glu Ala Ser Ala Ser Ala Asn Pro Ser Ile Asn 1 5 10 15

His Gln Pro Ser Thr Ile Thr Glu Ala Ala Arg Gln Arg Ile Leu Ile
20 25 30

Leu Asp Gly Ala Trp Gly Thr Gln Leu Gln Arg Ala Asn Leu Thr Glu
35 40 45

Ala Asp Phe Arg Trp Asp Glu Ala Asp Pro Thr Arg Met Tyr Arg Gly
50 55 60

Asn Phe Asp Leu Leu Gln Leu Thr Lys Pro Asp Val Ile Arg Ala Val 65 70 75 80

His Arg Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ala Ser Thr Asn Thr

- Phe Asn Ser Thr Thr Ile Ser Gln Ala Asp Tyr Gly Thr Glu Ala Leu 110

 Ala Tyr Ala Met Asn Arg Glu Gly Ala Arg Leu Ala Arg Glu Val Ala 115

 120

 125
- Asp Glu Phe Glu Ala Arg Asp Gly Lys Lys Arg Trp Val Ala Gly Ser 130 135 140
- Val Gly Pro Thr Asn Arg Thr Ala Thr Leu Ser Pro Asp Val Glu Arg 145 155 160
- Pro Glu Phe Arg Asn Val Thr Tyr Asp Asp Leu Val Ala Ala Tyr Ser 165 170 175
- Glu Ala Ile Thr Gly Leu Met Glu Gly Gly Ala Asp Leu Leu Ile 180 185 190
- Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Leu Phe Ala Ala 195 200 205
- Gln Asp Val Phe Ala Ala Gln Gly Arg Glu Leu Pro Val Met Leu Ser 210 215 220
- Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Pro 225 230 235 240
- Glu Ala Phe Ala Val Ser Thr Glu His Ala Gly Leu Phe Ser Leu Gly
 245 250 255
- Leu Asn Cys Ala Leu Gly Ala Asp Leu Leu Arg Pro His Leu Arg Ala 260 265 270
- Ile Ala Ala Asn Thr Glu Ala Leu Val Ser Val His Pro Asn Ala Gly 275 280 285
- Leu Pro Asn Ala Phe Gly Glu Tyr Asp Glu Thr Pro Glu His Thr Ala
 290 295 300
- Ala Val Leu Ala Asp Phe Ala Arg Glu Gly Leu Val Asn Ile Val Gly 305 310 315 320
- Gly Cys Cys Gly Thr Thr Pro Glu His Ile Lys Ala Ile Ala Glu Ala 325 330 335
- Val Lys Asp Ile Pro Pro Arg Gln Ala Leu Gln Leu Pro Pro Tyr Leu 340 345 350
- Arg Leu Ser Gly Leu Glu Ala Phe Thr Leu Thr Pro Glu Thr Asn Phe 355 360 365
- Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Pro Lys Phe Ser 370 380
- Lys Ala Ile Leu Ala Gly Asp Tyr Asp Ala Gly Leu Lys Ile Ala Arg 385 390 395 400
- Gln Gln Val Thr Asn Gly Ala Gln Ile Val Asp Ile Asn Phe Asp Glu 405 410 415

- Gly Met Leu Asp Gly Glu Gly Ala Met Val Lys Phe Leu Asn Leu Leu 420 425 430
- Ala Gly Glu Pro Asp Ile Ser Arg Val Pro Leu Met Leu Asp Ser Ser 435 440 445
- Lys Trp Glu Ile Leu Glu Ala Gly Leu Arg Arg Val Gln Gly Lys Ala
 450 460
- Val Val Asn Ser Ile Ser Leu Lys Asp Gly Glu Ala Arg Phe Leu Glu 465 470 475 480
- Arg Ala Arg Leu Leu Arg Arg Tyr Gly Ala Ala Ala Val Val Met Ala
 485 490 495
- Phe Asp Glu Gln Gly Gln Ala Asp Asn Leu Ala Arg Arg Glu Ile 500 510
- Leu Gly Arg Ala Tyr Arg Leu Leu Thr Glu Gln Ala Asp Phe Pro Pro 515 520 525
- Gln Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val Ala Thr Gly Ile 530 535 540
- Glu Glu His Asp Arg Tyr Ala Leu Asp Phe Ile Glu Ala Thr Arg Trp 545 550 555 560
- Ile Lys Glu Asn Leu Pro Ala Ala Lys Val Ser Gly Gly Ile Ser Asn 565 570 575
- Val Ser Phe Ser Phe Arg Gly Asn Asn His Val Arg Glu Ala Met His 580 585 590
- Ala Val Phe Leu Tyr His Ala Ile Arg Ala Gly Leu Asp Met Gly Ile
 595 600 605
- Val Asn Ala Gly Met Leu Ala Val Tyr Glu Asp Ile Glu Pro Glu Leu 610 620
- Arg Glu Ala Val Glu Asp Val Ile Leu Ala Arg Arg Pro Asp Ala Thr 625 630 635 640
- Glu Arg Leu Leu Thr Leu Ala Asp Arg Tyr Lys Asp Ile Lys Arg Glu 645 650 655
- Ser Ala Ala Gln Ser Ala Trp Arg Asp Leu Pro Val Gln Glu Arg Leu 660 665 670
- Arg His Ala Leu Val Gln Gly Val Ala Asp His Val Asp Glu Asp Ala 675 680 685
- Glu Ala Ala Tyr Gln Glu Leu Gly Ser Pro Leu Ala Val Ile Glu Gly 690 695 700
- Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly 705 710 715 720
- Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Lys 725 730 735
- Ala Val Ala Tyr Leu Thr Pro Tyr Leu Glu Ala Glu Lys Ala Glu Ser 740 745 750

- Ser Ser Lys Gly Lys Val Leu Leu Ala Thr Val Lys Gly Asp Val His 755 760 765
- Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Ala Cys Asn Gly Tyr 770 775 780
- Gln Val Thr Asp Leu Gly Val Met Val Pro Gly Glu Lys Ile Leu Asp
 785 790 795 800
- Glu Ala Glu Arg Leu Gly Ala Asp Val Ile Gly Leu Ser Gly Leu Ile 805 810 815
- Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Arg Glu Met Thr Arg 820 825 830
- Arg Gly Val Lys Thr Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg 835 840 845
- Ala His Thr Ala Val Lys Ile Asp Pro Ala Tyr Asp Gly Thr Val Val 850 855 860
- His Val Leu Asp Ala Ser Arg Ala Val Thr Val Thr Asn Asp Leu Leu 865 870 875 880
- Thr Asp Glu Ala Ala Tyr Ala Gly Arg Val Gln Gly Glu Tyr Asp Thr 885 890 895
- Leu Arg Glu Arg His Gly Glu Arg Gln Val Arg Leu Ile Ala Leu Ala 900 905 910
- Glu Ala Arg Ala Arg Ala Pro Gln Leu Ser Ala Ala Val Pro Pro Ala 915 920 925
- Pro His Asp Leu Gly Arg Gln Val Val Glu Gln Pro Ile Ala Glu Leu 930 935 940
- Leu Pro Phe Ile Asp Trp Thr Pro Phe Phe Ile Ala Trp Glu Met Lys 955 960
- Gly Ile Tyr Pro Gly Ile Leu Thr Asp Pro Leu Arg Gly Glu Glu Ala 965 970 975
- Arg Lys Leu Phe Ala Asp Ala Gln Ala Leu Leu Glu Gln Val Ile Ala 980 985 990
- Asp Gly Ser Leu Arg Ala Arg Gly Val Ile Gly Leu Trp Pro Ala His 995 1000 1005
- Gly Asp Asp Ile Val Leu Asp Asp Ala Ala Met Gly Arg Gly Glu Thr 1010 1015 1020
- Leu Asp Phe Glu Thr His Glu Leu Ala Ala Gly Arg Glu Pro Leu Pro 1025 1030 1035 1040
- Asn Met Pro Arg Leu His Thr Leu Arg Gln Gln Arg Asp Gln Thr Thr 1045 1050 1055
- Pro Asn Thr Ala Leu Ala Asp Phe Val Ala Glu Gly Gly Asp His Ile 1060 1065 1070
- Gly Ala Phe Ala Thr Ala Ile Phe Gly Ala Glu Glu Leu Ala Gln Gln

1075

1080

1085

Phe Glu Ala Gln His Asp Asp Tyr Asn Ser Ile Leu Val Lys Ala Val 1090 1095 1100

Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Lys Leu His Arg Asp Val 1105 1110 1115 1120

Arg Val Arg His Trp Gly Tyr Ala Glu Gly Glu Ala Leu Asp Asn Thr 1125 1130 1135

Asp Leu Ile Lys Glu Arg Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr 1140 1145 1150

Pro Ala Gln Pro Asp His Thr Glu Lys Arg Thr Leu Phe Glu Leu Leu 1155 1160 1165

Asp Ala Glu Ser Ile Gly Leu Arg Leu Thr Glu Ser Cys Ala Met Thr 1170 1175 1180

Pro Ala Ala Ala Val Ser Gly Leu Tyr Phe Ala His Pro Glu Ala Arg 1185 1190 1195 1200

Tyr Phe Ala Val Gly Arg Ile Gly Arg Asp Gln Val Glu Asn Tyr Ala 1205 1210 1215

Ala Arg Lys Gly Trp Thr Val Gln Glu Ala Glu Arg Trp Leu Gly Pro 1220 1225 1230

Leu Leu Ala Tyr Ser Ala Gly Pro Gly Pro Glu Ala Ser Gln Lys Ala 1235 1240 1245

Leu Gly Ala Glu Leu Thr Gly Ala Gln Ser 1250 1255

<210> 37

<211> 3642

<212> DNA

<213> Clostridium acetobutylicum

<220>

<221> CDS

<222> (1) .. (3639)

<223> RCA01265

<400> 37

ctt atg aat tct tca cta aag aat ttg tta aat aac aaa att tta gtt 48 Leu Met Asn Ser Ser Leu Lys Asn Leu Leu Asn Asn Lys Ile Leu Val 1 5 10 15

tta gat ggt gct atg gga aca tgt att caa tcc ttt aat cta gat gaa 96 Leu Asp Gly Ala Met Gly Thr Cys Ile Gln Ser Phe Asn Leu Asp Glu 20 25 30

ggc gac ttt aaa ggt tcc tta tct tgt aca tgt cat tcc aat caa aaa 144 Gly Asp Phe Lys Gly Ser Leu Ser Cys Thr Cys His Ser Asn Gln Lys 35 40 45

gga aac aat gat gtt tta aat tta acc aag cca gaa ata ata aaa gaa 192 Gly Asn Asn Asp Val Leu Asn Leu Thr Lys Pro Glu Ile Ile Lys Glu 50

6	5 ,	բ ոչ	'S AI	у ту	7 Let)	1 Ala	a Gly	y Ala	a As _j 7:	p Ile 5	e Ile	e Gl	u Th	a aac r Asn 80	240
111	L PII	e As	п Ат	a Th:	r Giv	ı Ile	Ser	Glr	90 90	a Asj	р Туз	: Ası	ı Mei	t Gl: 9!		288
aa. Ly	a at	a ta e Ty	t ga r As 10	БП	t aat e Asn	ttt Phe	aag Lys	999 Gly 105	/ Ala	aaa Lys	a cto s Lev	gca Ala	aaq Lyi	3 Gl	a gct ı Ala	336
tg: Cy:	t act	t ta r Ty 11	r ly	c aca r Thi	a aaa C Lys	cta Leu	aat Asn 120	Pro	aat Asn	aag Lys	g cct B Pro	aga Arg 125	, Ph€	gct Ala	gct Ala	384
ggt Gly	tci Sei 130	LIT	t ggg	g cct y Pro	aca Thr	aat Asn 135	Arg	act Thr	gct Ala	Sex	cta Leu 140	Sex	cca Pro	gat Asp	gtt Val	432
gaa Glu 145	ABI	cet Pro	t gg(c Gl _j	t ttt / Phe	aga Arg 150	aat Asn	gta Val	acc Thr	ttt Phe	gat Asp 155	Glu	cta Leu	tgt Cys	aat Asn	gcc Ala 160	480
tat Tyr	aaa Lys	cat His	caa Glr	ata 11e 165	GIU	gct Ala	cta Leu	ata Ile	gat Asp 170	gga Gly	ggt	gta Val	gac	Leu 175	Leu	528
tta Leu	att Ile	gaa Glu	act Thr 180	ata Ile	ttt Phe	gat Asp	act Thr	tta Leu 185	aac Asn	gct Ala	aga Arg	gca Ala	gca Ala 190	Ile	ttt Phe	576
gca Ala	gca Ala	gaa Glu 195	Inr	gta Val	ttt Phe	gaa Glu	aat Asn 200	aaa Lys	aaa Lys	ata Ile	aaa Lys	ctt Leu 205	cct Pro	att Ile	ata Ile	624
att Ile	tca Ser 210	GIA	aca Thr	ata Ile	gct Ala	gat Asp 215	aaa Lys	agt Ser	gga Gly	aga Arg	ata Ile 220	tta Leu	tcc Ser	ggt Gly	caa Gln	672
act Thr 225	ctt Leu	gac Asp	gct Ala	ttt Phe	gca Ala 230	gaa Glu	agt Ser	tta Leu	aaa Lys	aac Asn 235	gaa Glu	aat Asn	ata Ile	att Ile	gct Ala 240	720
ata Ile	Gly 999	ctt Leu	aat Asn	tgt Cys 245	tcc Ser	ttt Phe	ggt Gly	gct Ala	gaa Glu 250	gaa Glu	ctt Leu	ata Ile	cct Pro	ttt Phe 255	ata Ile	768
aaa Lys	aga Arg	ctc Leu	tct Ser 260	gaa Glu	aca Thr	caa Gln i	Asn .	aga Arg 265	tat Tyr	ata Ile	tcc Ser	ttt Phe	cat His 270	cca Pro	aac Asn	816
gca Ala	gga Gly	ctt Leu 275	cca Pro	aac Asn	tcc (Ser 1	Leu (ggt (Gly (280	gaa Glu	tat Tyr	gaa Glu	Glu	ctg Leu 285	cca Pro	gag Glu	gaa Glu	864
TILL	gct Ala 290	agc Ser	att Ile	gta Val	råe 1	aaa t Lys 1 295	tta q Leu i	gca Ala	ctt Leu	Glu	gga Gly : 300	cat His	tta Leu	aat Asn	ata Ile	912
gtt	gga	ggc	tgc	tgt (ggc a	act a	aca d	ca	gaa (cat	ata .	aga 🤈	gca	ata	agc	960

WO 03/087386 PCT/EP03/04010 163 Val Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Arg Ala Ile Ser 310 315 age gta gtt aaa gge att tet eea aga aaa gtt eea aac ttg gaa eee 1008 Ser Val Val Lys Gly Ile Ser Pro Arg Lys Val Pro Asn Leu Glu Pro 325 aaa aca att tac agc gga cta gaa aac ata aaa att gat aag aac agt 1056 Lys Thr Ile Tyr Ser Gly Leu Glu Asn Ile Lys Ile Asp Lys Asn Ser 340 345 aac ttc ata aat ata ggc gaa aga aca aat gta gcg ggc tca aga aaa 1104 Asn Phe Ile Asn Ile Gly Glu Arg Thr Asn Val Ala Gly Ser Arg Lys 355 ttc gca agg ctt ata cgt gaa aaa aat tat gag gag gct cta acc att 1152 Phe Ala Arg Leu Ile Arg Glu Lys Asn Tyr Glu Glu Ala Leu Thr Ile 370 375 gca aga cat cag gtt gaa aat ggt gcc caa att ata gat ata aat ttt 1200 Ala Arg His Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile Asn Phe 385 395 gat gat gca ctt tta gat gct cgc tct gaa atg gaa aca ttt tta aga Asp Asp Ala Leu Leu Asp Ala Arg Ser Glu Met Glu Thr Phe Leu Arg 405 410 ctt att gca agt gaa cct gaa ata tca aaa gtt cca gtt atg ata gac 1296 Leu Ile Ala Ser Glu Pro Glu Ile Ser Lys Val Pro Val Met Ile Asp 425 tcc tct aat ttt gaa gtt tta aaa gtt gga tta aag tct att caa ggt Ser Ser Asn Phe Glu Val Leu Lys Val Gly Leu Lys Ser Ile Gln Gly 440 aaa gcc ata gta aat tcc ata agt ctt aag gtt gga gaa gaa aag ttc Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Val Gly Glu Glu Lys Phe 455 att gaa gag gca aaa ttt ata aag aac ttt ggc gct ggc gta gtt gta Ile Glu Glu Ala Lys Phe Ile Lys Asn Phe Gly Ala Gly Val Val 470 atg gcc ttt gac gaa ggt caa gca gct act tat gaa aga aaa att Met Ala Phe Asp Glu Glu Gly Gln Ala Ala Thr Tyr Glu Arg Lys Ile 490 gaa atc tgc aag aga gct tat act att ctc aca gaa aaa gtt gag ttt 1536 Glu Ile Cys Lys Arg Ala Tyr Thr Ile Leu Thr Glu Lys Val Glu Phe 500 505 cca cct gaa aat ata ata ttt gat cca aat ata cta tct ata gcg aca 1584 Pro Pro Glu Asn Ile Ile Phe Asp Pro Asn Ile Leu Ser Ile Ala Thr 515 520 gga att gaa gaa cat gac aac tat gca gtt aat tac ata aaa gct gtt 1632 Gly Ile Glu Glu His Asp Asn Tyr Ala Val Asn Tyr Ile Lys Ala Val 530 535

aaa tgg ata aaa gag aat cta cca tac gct aaa gtc agc ggt gga gtt . Lys Trp Ile Lys Glu Asn Leu Pro Tyr Ala Lys Val Ser Gly Gly Val

555

550

age Sei	c aad r Ası	c ct	c tc	r Pho	e Se	t tti r Phe	t ago	9 99°	2 aa1 7 As1 570	a Ası	c gc	a ata a Ile	a aga Arg	a aga g Arg 57!	a gct g Ala 5	1728
atg Met	g cat	t to	t gti r Vai 580	l Phe	c cti	tac 1 Tyr	cat His	gca 8 Ala 585	lle	a aac B Asi	get Ala	t gga a Gly	ate Met 590	gaA :	atg Met	1776
ggt Gly	att Ile	gt(Va) 599	L Ası	cca Pro	a gca o Ala	atg Met	att : Ile 600	: Asr	i tta	a tat 1 Tyr	: Aaj	gat Asp 605	ı I I e	gat Asp	aag Lys	1824
gat	Ctt Lev 610	Let	gaa Glu	a aag 1 Lys	gtt Val	gag Glu 615	Asn	gtt Val	gta Val	t cta	aat 1 Asi 620	ı Lys	tca Ser	tct Ser	aac Asn	1872
gct Ala 625	ser	gaa Glu	tca Ser	tta Leu	cta Leu 630	Glu	ttt Phe	gct Ala	cas Glr	acg Thr 635	Ty	aaa Lys	aag Lys	acg Thr	act Thr 640	1920
GIU.	Thr	Leu	. Glu	645	His	Glu	Asp	Glu	Trp 650	Arg	Glr	Lys	Ser	Pro 655		1968
GIU	Arg	Leu	660	Tyr	Ala	Leu	Val	Lys 665	Gly	Asn	Val	gaa Glu	Phe 670	Ile	Glu	2016
Glu	Asp	11e 675	Glu	Glu	Ala	Arg	Lys 680	Glu	Tyr	Thr	naA	gca Ala 685	Leu	Glu	Ile	2064
Ile	690	Val	Pro	Leu	Met	Asn 695	Gly	Met	Ľув	Lys	Val 700		Lys	Leu	Phe	2112
705	Glu	Gly	Lys	Met	Phe 710	Leu	Pro	Gln	Val	Val 715	Lys	agt Ser	Ala	Arg	Val 720	2160
Met	Lys	Lys	Ala	Val 725	Glu	Сув	Leu	Leu	Pro 730	Tyr	Ile	aac Asn	Glu	Glu 735	Lys	2208
ser	Lys	Asn	740	Asn	Lys	Ser	Ala	Gly 745	Lys	Val	Val	ttt Phe	Ala 750	Thr	Val	2256
гàв	GIA	755	Val	His	Asp	Ile	Gly 760	Lys	Aen	Ile	Val	tct Ser 765	Val	Val	Leu	2304
ser	770	Asn	Asn	Phe	Glu	Val 775	Ile	Asp	Leu	Gly	Val 780	atg Met	Val	Pro	Pro	2352
785	Thr	116	Leu	Glu	Thr 790	Ala	Lys	Arg	Glu	Asn 795	Ala	gat Asp	Ile	Ile	Ala 800	2400
tta	agt	ggt	tta	att	aca	cct	tct	ctt	aat	gaa	atg	gct	tat	gta	gct	2448

• .	_		_		_					105						
Leu	Sei	r Gl	y Lev	805	e Thr	Pro	Ser	Leu	810		Met	: Ala	Туг	Val 815	Ala	
gaa Glu	gaa Glu	a atq 1 Mei	aaa Lys 820	Arg	g ctt g Leu	aat Asn	ttt Phe	gat Asp 825	Ile	cca Pro	ctt Leu	atg Met	gtg Val 830	Gly	ggt Gly	2496
Ala	Ala	835	s Ser	Lys	Thr	His	Thr 840	Ala	Leu	Lys	Leu	Ala B45	Thr	Lys	tat Tyr	2544
Lys	Tyr 850	· Va]		His	Ser	Thr 855	Asp	Ala	Ser	, yab	Ala 860	Val	Thr	Val	Ala	2592
Lув 865	Asn	Lev	atg Met	Ser	870	Asn	Lys	Phe	Thr	Phe 875	Leu	Glu	Lys	Leu	Asn 880	2640
Glu	Glu	Tyr	tct Ser	Lys 885	Ile	Arg	Glu	Thr	Phe 890	Ser	Thr	Asn	Lys	11e 895	Glu	2688
Leu	Ile	Ser	lle 900	Gln	Asn	Ala	Arg	Ъув 905	Asn	Arg	Phe	Thr	Ile 910	Asp	Trp	2736
Asn	Lys	7hr 915		Ile	Thr	Glu	Pro 920	Lув	Phe	Val	Gly	11e 925	Lys	Lys	Leu	2784
Gln	A1a 930	Val	cct Pro	Ile	Asn	Glu 935	Leu	Arg	ГÀв	Tyr	Ile 940	Asp	Trp	Thr	Phe	2832
945	Pne	Thr	tct Ser	Trp	950	Met	Gly	Met	Asn	Tyr 955	Pro	Lys	Ile	Met	Lys 960	2880
ABD	Pro	Lys	tac Tyr	965	Ala	Glu	Ala	Gln	Lys 970	Leu	Phe	Lys	Asp	Ala 975	Asn	2928
Glu	Met	Leu	gat Asp 980	Leu	Leu	Gln	Lys	Glu 985	Asn	Leu	Ile	Thr	Сув 990	Asn	Gly	2976
Val	Phe	Gly 995	ata Ile	Phe	Pro	Ala 1	Asn .000	Ser	Val	Aan	Aap I	Asp 005	Ile	Glu	Ile	3024
Tyr 1	Thr 010	Asp	aaa Lys	Gly	Thr . 1	Val 015	Thr	Ile	Asn	Thr 1	Leu 020	Arg	Gln	Gln	Gln	3072
ata Ile 1025	ctt Leu	aaa Lys	gac Asp	Ser	gat Asp 1030	tat Tyr	aaa Lys	gct Ala	cta Leu	tct Ser 1035	qaA	tat Tyr	atc Ile	gct Ala	cca Pro 1040	3120
aag (Lys (ggt Gly	att Ile	ggc Gly 1	atc Ile 045	aaa Lys .	gat Asp	tat Tyr	Ile	ggt Gly 050	ggt Gly	ttt Phe	att Ile	Val	act Thr 055	gct Ala	3168

100	
gga ata ggt gca aag gaa tat tcc gat aaa tta aag aaa aaa tgc gac Gly Ile Gly Ala Lys Glu Tyr Ser Asp Lys Leu Lys Lys Lys Cys Asp 1060 1065 1070	3216
gat tat gga gct act atg ctt aaa ctt ata tgc gat aga ctt gca gag Asp Tyr Gly Ala Thr Met Leu Lys Leu Ile Cys Asp Arg Leu Ala Glu 1075 1080 1085	3264
gcc ttt tca gaa ctt ctt cac cta agg gta aga aaa gaa tac tgg gga Ala Phe Ser Glu Leu Leu His Leu Arg Val Arg Lys Glu Tyr Trp Gly 1090 1095 1100	3312
tac tct caa gat gaa aac tta tcc tta gaa aaa ctt ctt aaa gga agt Tyr Ser Gln Asp Glu Asn Leu Ser Leu Glu Lys Leu Leu Lys Gly Ser 1105 1110 1115 1120	3360
tac aga ggg ata aaa cca gct att gga tat cct tct att ccc gat cac Tyr Arg Gly Ile Lys Pro Ala Ile Gly Tyr Pro Ser Ile Pro Asp His 1125 1130 1135	3408
tct gaa aaa gca aag tta ttt gat tta ctt tta ggt aaa act tca ata Ser Glu Lys Ala Lys Leu Phe Asp Leu Leu Leu Gly Lys Thr Ser Ile 1140 1145 1150	3456
gga gtg gaa ttg acg gaa agt tat atg atg aat cca act tca agt gta Gly Val Glu Leu Thr Glu Ser Tyr Met Met Asn Pro Thr Ser Ser Val 1155 1160 1165	3504
tgc ggt ttg tat ttt gca aat gaa cga gca aaa tac ttt aat ata aat : Cys Gly Leu Tyr Phe Ala Asn Glu Arg Ala Lys Tyr Phe Asn Ile Asn 1170 1180	3552
aaa ata gga aaa gat caa ctt gag gac tat gct gtt cga agt aat aaa : Lys Ile Gly Lys Asp Gln Leu Glu Asp Tyr Ala Val Arg Ser Asn Lys 1185 1190 1195 1200	3600
gac att aat gaa ata aaa aaa tta tta gat act ctg tta taa Asp Ile Asn Glu Ile Lys Lys Leu Leu Asp Thr Leu Leu 1205 1210	3642
<210> 38 <211> 1213 <212> PRT <213> Clostridium acetobutylicum	
<400> 38	
Leu Met Asn Ser Ser Leu Lys Asn Leu Leu Asn Asn Lys Ile Leu Val 1 5 10 15	
Leu Asp Gly Ala Met Gly Thr Cys Ile Gln Ser Phe Asn Leu Asp Glu 20 25 30	
Gly Asp Phe Lys Gly Ser Leu Ser Cys Thr Cys His Ser Asn Gln Lys 35 40 45	
Gly Asn Asn Asp Val Leu Asn Leu Thr Lys Pro Glu Ile Ile Lys Glu 50 55 60	
Ile His Lys Arg Tyr Leu Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn 65 70 75 80	

- Thr Phe Asn Ala Thr Glu Ile Ser Gln Lys Asp Tyr Asn Met Gln Asp 85 90 95
- Lys Ile Tyr Asp Ile Asn Phe Lys Gly Ala Lys Leu Ala Lys Glu Ala 100 105 110
- Cys Thr Tyr Tyr Thr Lys Leu Asn Pro Asn Lys Pro Arg Phe Ala Ala 115 120 125
- Gly Ser Ile Gly Pro Thr Asn Arg Thr Ala Ser Leu Ser Pro Asp Val
- Glu Asn Pro Gly Phe Arg Asn Val Thr Phe Asp Glu Leu Cys Asn Ala 145 150 155 160
- Tyr Lys His Gln Ile Glu Ala Leu Ile Asp Gly Gly Val Asp Leu Leu 165 170 175
- Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Arg Ala Ala Ile Phe 180 185 190
- Ala Ala Glu Thr Val Phe Glu Asn Lys Lys Ile Lys Leu Pro Ile Ile 195 200 205
- Ile Ser Gly Thr Ile Ala Asp Lys Ser Gly Arg Ile Leu Ser Gly Gln 210 215 220
- Thr Leu Asp Ala Phe Ala Glu Ser Leu Lys Asn Glu Asn Ile Ile Ala 225 230 235 240
- Ile Gly Leu Asn Cys Ser Phe Gly Ala Glu Glu Leu Ile Pro Phe Ile 245 250 255
- Lys Arg Leu Ser Glu Thr Gln Asn Arg Tyr Ile Ser Phe His Pro Asn 260 265 270
- Ala Gly Leu Pro Asn Ser Leu Gly Glu Tyr Glu Glu Leu Pro Glu Glu 275 280 285
- Thr Ala Ser Ile Val Lys Lys Leu Ala Leu Glu Gly His Leu Asn Ile 290 295 300
- Val Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Arg Ala Ile Ser 305 310 315 320
- Ser Val Val Lys Gly Ile Ser Pro Arg Lys Val Pro Asn Leu Glu Pro 325 330 335
- Lys Thr Ile Tyr Ser Gly Leu Glu Asn Ile Lys Ile Asp Lys Asn Ser 340 345 350
- Asn Phe Ile Asn Ile Gly Glu Arg Thr Asn Val Ala Gly Ser Arg Lys 355 360 365
- Phe Ala Arg Leu Ile Arg Glu Lys Asn Tyr Glu Glu Ala Leu Thr Ile 370 375 380
- Ala Arg His Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile Asn Phe 385 390 395 400
- Asp Asp Ala Leu Leu Asp Ala Arg Ser Glu Met Glu Thr Phe Leu Arg

410

405

- Leu Ile Ala Ser Glu Pro Glu Ile Ser Lys Val Pro Val Met Ile Asp 420 425 430
- Ser Ser Asn Phe Glu Val Leu Lys Val Gly Leu Lys Ser Ile Gln Gly
 435 440 445
- Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Val Gly Glu Glu Lys Phe
 450 460
- Ile Glu Glu Ala Lys Phe Ile Lys Asn Phe Gly Ala Gly Val Val 480
- Met Ala Phe Asp Glu Glu Gly Gln Ala Ala Thr Tyr Glu Arg Lys Ile
 485 490 495
- Glu Ile Cys Lys Arg Ala Tyr Thr Ile Leu Thr Glu Lys Val Glu Phe 500 505 510
- Pro Pro Glu Asn Ile Ile Phe Asp Pro Asn Ile Leu Ser Ile Ala Thr 515 525
- Gly Ile Glu Glu His Asp Asn Tyr Ala Val Asn Tyr Ile Lys Ala Val 530 540
- Lys Trp Ile Lys Glu Asn Leu Pro Tyr Ala Lys Val Ser Gly Gly Val 545 550 555 560
- Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Asp Ala Ile Arg Arg Ala 565 570 575
- Met His Ser Val Phe Leu Tyr His Ala Ile Asn Ala Gly Met Asp Met 580 590
- Gly Ile Val Asn Pro Ala Met Ile Asp Leu Tyr Asp Asp Ile Asp Lys 595 600 605
- Asp Leu Leu Glu Lys Val Glu Asn Val Val Leu Asn Lys Ser Ser Asn 610 620
- Ala Ser Glu Ser Leu Leu Glu Phe Ala Gln Thr Tyr Lys Lys Thr Thr 625 635 640
- Glu Thr Leu Glu Lys His Glu Asp Glu Trp Arg Gln Lys Ser Pro Ser 645 650 655
- Glu Arg Leu Ser Tyr Ala Leu Val Lys Gly Asn Val Glu Phe Ile Glu
 660 665 670
- Glu Asp Ile Glu Glu Ala Arg Lys Glu Tyr Thr Asn Ala Leu Glu Ile 675 680 685
- Ile Glu Val Pro Leu Met Asn Gly Met Lys Lys Val Gly Lys Leu Phe 690 695 700
- Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val 705 710 715 720
- Met Lys Lys Ala Val Glu Cys Leu Leu Pro Tyr Ile Asn Glu Glu Lys
 725 730 735

169

Ser Lys Asn His Asn Lys Ser Ala Gly Lys Val Val Phe Ala Thr Val 740 745 750

- Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser Val Val Leu 755 760 765
- Ser Cys Asn Asn Phe Glu Val Ile Asp Leu Gly Val Met Val Pro Pro 770 780
- Glu Thr Ile Leu Glu Thr Ala Lys Arg Glu Asn Ala Asp Ile Ile Ala 785 790 795 800
- Leu Ser Gly Leu Ile Thr Pro Ser Leu Asn Glu Met Ala Tyr Val Ala 805 810 815
- Glu Glu Met Lys Arg Leu Asn Phe Asp Ile Pro Leu Met Val Gly Gly 820 825 830
- Ala Ala Thr Ser Lys Thr His Thr Ala Leu Lys Leu Ala Thr Lys Tyr 835 840 845
- Lys Tyr Val Val His Ser Thr Asp Ala Ser Asp Ala Val Thr Val Ala 850 855 860
- Lys Asn Leu Met Ser Glu Asn Lys Phe Thr Phe Leu Glu Lys Leu Asn 865 870 875 880
- Glu Glu Tyr Ser Lys Ile Arg Glu Thr Phe Ser Thr Asn Lys Ile Glu 885 890 895
- Leu Ile Ser Ile Gln Asn Ala Arg Lys Asn Arg Phe Thr Ile Asp Trp
 900 905 910
- Asn Lys Thr Lys Ile Thr Glu Pro Lys Phe Val Gly Ile Lys Lys Leu 915 920 925
- Gln Ala Val Pro Ile Asn Glu Leu Arg Lys Tyr Ile Asp Trp Thr Phe 930 935 940
- Phe Phe Thr Ser Trp Asp Met Gly Met Asn Tyr Pro Lys Ile Met Lys 945 950 955 960
- Asp Pro Lys Tyr Gly Ala Glu Ala Gln Lys Leu Phe Lys Asp Ala Asn 965 970 975
- Glu Met Leu Asp Leu Leu Gln Lys Glu Asn Leu Ile Thr Cys Asn Gly 980 985 990
- Val Phe Gly Ile Phe Pro Ala Asn Ser Val Asn Asp Asp Ile Glu Ile 995 1000 1005
- Tyr Thr Asp Lys Gly Thr Val Thr Ile Asn Thr Leu Arg Gln Gln Gln 1010 1015 1020
- Ile Leu Lys Asp Ser Asp Tyr Lys Ala Leu Ser Asp Tyr Ile Ala Pro 1025 1030 1035 1040
- Lys Gly Ile Gly Ile Lys Asp Tyr Ile Gly Gly Phe Ile Val Thr Ala 1045 1050 1055
- Gly Ile Gly Ala Lys Glu Tyr Ser Asp Lys Leu Lys Lys Lys Cys Asp 1060 1065 1070

									•	10					
Asp '	Tyr 1	Gly 075	Ala	Thr	Met	Leu	Lys 1080	Leu	Ile	Су́в		Arg 1085	Leu	Ala	Glu
Ala - 1	Phe :	Ser	Glu	Leu	Leu	His 1095	Leu	Arg	Val		Lув 1100	Glu	Tyr	Trp	Gly
Tyr :	Ser (Gln	Asp	Glu 1	Asn 110	Leu	Ser	Leu		Lys 1115	Leu	Leu	Lys	-	Ser

Tyr Arg Gly Ile Lys Pro Ala Ile Gly Tyr Pro Ser Ile Pro Asp His 1125 1130

Ser Glu Lys Ala Lys Leu Phe Asp Leu Leu Leu Gly Lys Thr Ser Ile 1145

Gly Val Glu Leu Thr Glu Ser Tyr Met Met Asn Pro Thr Ser Ser Val 1155 1160

Cys Gly Leu Tyr Phe Ala Asn Glu Arg Ala Lys Tyr Phe Asn Ile Asn 1175

Lys Ile Gly Lys Asp Gln Leu Glu Asp Tyr Ala Val Arg Ser Asn Lys 1190 1195 1200

Asp Ile Asn Glu Ile Lys Lys Leu Leu Asp Thr Leu Leu

<210> 39 <211> 3954 <212> DNA <213> Caulobacter_crescentus <220> <221> CDS <222> (1)..(3951) <223> RCO02271 <400> 39

- atg acc gat etc tee ate ege gee aac ege gte gee etg aag gee 48 Met Thr Asp Leu Ser Ile Arg Ala Asn Arg Val Ala Ala Leu Lys Ala
- gee gee aag gag egt att ete att ete gae gge tee tgg gge gtg atg 96 Ala Ala Lys Glu Arg Ile Leu Ile Leu Asp Gly Ser Trp Gly Val Met
- ttc cag aag aag ggg ctg acc gag gcc gac tac cgc gcc gag cgc ttc Phe Gln Lys Lys Gly Leu Thr Glu Ala Asp Tyr Arg Ala Glu Arg Phe
- ged ged tad aad ggd dag atg aag ggd aat aad gad atd ctg tgd ctg 192 Ala Ala Tyr Asn Gly Gln Met Lys Gly Asn Asn Asp Ile Leu Cys Leu 55
- acg cgg ccc gat ctc gtg gcc gag ctg cac gac gcc tat ttc agc gcc 240 Thr Arg Pro Asp Leu Val Ala Glu Leu His Asp Ala Tyr Phe Ser Ala 75
- ggc gcc gac atc tcc gag acc aac acc ttc tcg ggc acc acc atc gcc Gly Ala Asp Ile Ser Glu Thr Asn Thr Phe Ser Gly Thr Thr Ile Ala

90

85 95 cag gcc gac tat cat ctg ggt gaa cag gat gtc tgg gac atc aac ctg Gln Ala Asp Tyr His Leu Gly Glu Gln Asp Val Trp Asp Ile Asn Leu 105 gaa ggc gcc aag atc ggc cgc tcg gtg gcc gac cgc tgg aac gcg cag 384 Glu Gly Ala Lys Ile Gly Arg Ser Val Ala Asp Arg Trp Asn Ala Gln aat eee gae ege eeg aag tte ate gee gge teg atg ggg eeg etg aac 432 Asn Pro Asp Arg Pro Lys Phe Ile Ala Gly Ser Met Gly Pro Leu Asn 135 gtc atg ctg tcg tcg tcg gac gtg aac gat ccg ggc gcg cgc aag 480 Val Met Leu Ser Met Ser Ser Asp Val Asn Asp Pro Gly Ala Arg Lys 150 gtg acc ttc gac cag gtc tac gag gcc tat cgc cag cag gtg gat gcg 528 Val Thr Phe Asp Gln Val Tyr Glu Ala Tyr Arg Gln Gln Val Asp Ala ctt tac cag ggc ggg gtc gat ctc ttc ctg atc gag acc atc acc gac 576 Leu Tyr Gln Gly Gly Val Asp Leu Phe Leu Ile Glu Thr Ile Thr Asp 180 185 acc ctg aac tgc aag gcc gcg atc aag gcg atc ctg gac tgg cgc gac 624 Thr Leu Asn Cys Lys Ala Ala Ile Lys Ala Ile Leu Asp Trp Arg Asp 195 gag ggc cac gag gag ctg ccg atc tgg atc agc ggc acc atc acc gat 672 Glu Gly His Glu Glu Leu Pro Ile Trp Ile Ser Gly Thr Ile Thr Asp 210 cgc tcg ggc cgc acc ctg tcg ggc cag acg gcc gag gcg ttc tgg aac 720 Arg Ser Gly Arg Thr Leu Ser Gly Gln Thr Ala Glu Ala Phe Trp Asn 225 235 age gte aag cae gee aag eeg tte gea gtg gge tte aac tge gee etg 768 Ser Val Lys His Ala Lys Pro Phe Ala Val Gly Phe Asn Cys Ala Leu 245 ggc gcg gat ttg atg cgt ccg cac atc gcc gag atg gcc cgt atc gcc 816 Gly Ala Asp Leu Met Arg Pro His Ile Ala Glu Met Ala Arg Ile Ala gac acc ctg gtc gca gcc tat ccc aac gcc ggc ctg ccc aac gcc atg Asp Thr Leu Val Ala Ala Tyr Pro Asn Ala Gly Leu Pro Asn Ala Met ggc cag tac gac gag gag ccg cac gag acc ggc cac gcc ctg cac gag 912 Gly Gln Tyr Asp Glu Glu Pro His Glu Thr Gly His Ala Leu His Glu 295 tgg gcc aag gac ggc ctc gtc aac atc ctg ggc ggc tgc tgc ggc acg 960 Trp Ala Lys Asp Gly Leu Val Asn Ile Leu Gly Gly Cys Cys Gly Thr 315 aca ccg gac cac atc cgt cac gtc gcc gac gag gtg cgc ggc gtg acg 1008 Thr Pro Asp His Ile Arg His Val Ala Asp Glu Val Arg Gly Val Thr 325 330 335

										72						
ccg Pro	cgc Arg	cag Gln	atc Ile 340	Pro	gag Glu	cgc	Pro	aag Lys 345	gcc Ala	atg Met	cgc	ctg Leu	gcg Ala 350	ggc	ctc Leu	1056
gaa Glu	Pro	ttc Phe 355	gag Glu	ttg Leu	gct Ala	tag Xaa	tgg Trp 360	cta Leu	cgg Arg	ccg Pro	caa Gln	att Ile 365	ccc Pro	ttc Phe	tcc Ser	1104
cct Pro	tgc Cys 370	ggg Gly	aga Arg	agg Arg	tgt Cys	cgc Arg 375	cga Arg	agg Arg	cga Arg	cgg Arg	atg Met 380	agg Arg	ggt Gly	ctc Leu	gcc Ala	1152
ggc Gly 385	cct Pro	tca Ser	acc Thr	gct Ala	gtc Val 390	tcg Ser	cgg Arg	cgg Arg	cga Arg	cgt Arg 395	tct Ser	tca Ser	acc Thr	cct Pro	cat His 400	1200
ccg Pro	acc Thr	cgc Arg	tgc Cys	gcg Ala 405	ggc Gly	cac His	ctt Leu	ctc Leu	ccg Pro 410	caa Gln	99 9 Gly	gag Glu	aag Lys	gga Gly 415	tga Xaa	1248
ctg Leu	cta Leu	ttg Leu	gat Asp 420	cct Pro	gaa Glu	atg Met	cgc Arg	ccc Pro 425	gtc Val	ttc Phe	gtc Val	aac Asn	atc Ile 430	ggt Gly	gag Glu	1296
cgc Arg	acc Thr	aac Asn 435	gtc Val	acc Thr	ggc	tcg Ser	gcc Ala 440	aag Lys	ttc Phe	aag Lys	aag Lys	ctg Leu 445	atc Ile	gtc Val	gaa Glu	1344
Gly 999	aac Asn 450	tat Tyr	ccc Pro	gag Glu	gcg Ala	ctg Leu 455	tcg Ser	gtc Val	gcg Ala	cgc Arg	cag Gln 460	cag Gln	gtc Val	gag Glu	gcc Ala	1392
999 Gly 465	gcc Ala	cag Gln	gtc Val	atc Ile	gac Asp 470	gtg Val	aac Asn	atg Met	gac Asp	gag Glu 475	ggt Gly	ctg Leu	ctg Leu	gac Asp	agc Ser 480	1440
cag Gln	cag Gln	gcc Ala	atg Met	gtc Val 485	acc Thr	ttc Phe	ctg Leu	aat Asn	ctg Leu 490	atg Met	gcg Ala	gcc Ala	gag Glu	ccc Pro 495	gac Asp	1488
atc Ile	gcg Ala	cgc Arg	gtg Val 500	ccg Pro	gtg Val	atg Met	atc Ile	gac Asp 505	agc Ser	tcc Ser	aag Lys	tgg Trp	gag Glu 510	gtg Val	atc Ile	1536
gag Glu	gcg Ala	ggc Gly 515	ctg Leu	aag Lys	tgc Cys	gta Val	caa Gln 520	ggc Gly	aag Lys	gcg Ala	atc Ile	gtc Val 525	aac Asn	tcg Ser	atc Ile	1584
agc Ser	ctg Leu 530	aag Lys	gaa Glu	ggc Gly	gag Glu	gaa Glu 535	aag Lys	ttc Phe	ctc Leu	gaa Glu	cag Gln 540	gcc Ala	acg Thr	ctc Leu	tgc Cys	1632
ctg Leu 545	cgc Arg	tat Tyr	ggc Gly	gca Ala	gcc Ala 550	gtg Val	gtg Val	gtc Val	atg Met	gcc Ala 555	ttc Phe	gac Asp	gag Glu	gtt Val	ggc Gly 560	1680
cag Gln	gcc Ala	gac Asp	acc Thr	gaa Glu 565	aag Lys	cgc Arg	aag Lys	gtc Val	gag Glu 570	atc Ile	tgt Cys	acg Thr	cgg Arg	gcc Ala 575	tac Tyr	1728
aac Asn	acg Thr	ctc Leu	gtg Val	gac Asp	aag Lys	gtc Val	ggc Gly	ttc Phe	ccg Pro	ccc Pro	gag Glu	gac Asp	atc Ile	atc Ile	ttc Phe	1776

	580		585	173		590	
gac ccc aac Asp Pro Asn 595	atc ttc Ile Phe	gcc gtg Ala Val	gcg acg Ala Thr 600	ggg atc Gly Ile	gag gag Glu Glu 605	cac gac His Asp	aac 1824 Asn
tac gcc gtc Tyr Ala Val 610	gac ttc Asp Phe	atc gag Ile Glu 615	gcc acg Ala Thr	cgg cgc Arg Arg	atc aag Ile Lys 620	cag atg Gln Met	ttg 1872 Leu
ccc tat gcg Pro Tyr Ala 625	cgg gtg Arg Val	tcg ggc Ser Gly 630	ggg gtg Gly Val	tcg aac Ser Asn 635	gtc tcg Val Ser	ttc agc Phe Ser	ttc 1920 Phe 640
cgg ggc aat Arg Gly Asn	gag ccg Glu Pro 645	gtg cgc Val Arg	cgg gcg Arg Ala	atc cac Ile His 650	tcg gtg Ser Val	ttc ctg Phe Leu 655	tac 1968 Tyr
cac gcc atc His Ala Ile	aac gcc Asn Ala 660	ggc atg Gly Met	gac atg Asp Met 665	ggc atc Gly Ile	gtc aac Val Asn	gcc ggc Ala Gly 670	gac 2016 Asp
ctg ccg gtc Leu Pro Val 675	tat gac Tyr Asp	gac atc Asp Ile	gat ccg Asp Pro 680	gcc ctg Ala Leu	cgc gag Arg Glu 685	gcc gtc Ala Val	gag 2064 Glu
gac gtg atc Asp Val Ile 690	Leu Asn	Arg Pro 695	Gln Arg	Asp Pro	Val Met 700	Thr Asn	Thr
gag cgc ctg Glu Arg Leu 705	Val Glu	Met Ala 710	Pro Arg	Tyr Lys 715	Gly Glu	Lys Gly	Gln 720
cag cag gtc Gln Gln Val	Ala Asn 725	Leu Glu	Trp Arg	Lys Gly 730	Thr Val	Asn Glu 735	Arg
ctg acc cat Leu Thr His	gct ctc Ala Leu 740	gtt cac Val His	ggc atc Gly Ile 745	acc gag Thr Glu	ttc atc Phe Ile	gag cag Glu Gln 750	gac 2256 Asp
acc gag gag Thr Glu Glu 755	Ala Arg	Leu Ala	Ala Glu 760	Arg Pro	Leu His 765	Val Ile	Glu
ggc ccg ctg Gly Pro Leu 770	Met Asp	Gly Met 775	Asn Val	Val Gly	Asp Leu 780	Phe Gly	Ala
ggc aag atg Gly Lys Met 785	Phe Leu	ccc cag q Pro Gln \ 790	gtg gtg Val Val	aag tcg Lys Ser 795	gcc cgc Ala Arg	gtg atg Val Met	aag 2400 Lys 800
cag gcc gtc Gln Ala Val	gcc tgg (Ala Trp 1 805	ctg atg (Leu Met 1	ccg ttc Pro Phe	atg gag Met Glu 810	gcc gag Ala Glu	aag gaa Lys Glu 815	ggc 2448 Gly
cag gag cgc Gln Glu Arg	aag gcc g Lys Ala 1 820	gcc ggc a Ala Gly 1	aag gtg Lys Val 825	ctg atg Leu Met	gcc acc Ala Thr	gtc aag Val Lys 830	ggc 2496 Gly

										74 ·			•			
gac Asp	gtc Val	cac His 835	gac Asp	atc Ile	ggt Gly	aag Lys	aac Asn 840	atc Ile	gtc Val	ggc Gly	gtc Val	gtg Val 845	ctg Leu	cag Gln	tgt Cys	2544
aac Asn	aac Asn 850	tac Tyr	gag Glu	gtc Val	gtg Val	gac Asp 855	ctg Leu	ggt Gly	gtc Val	atg Met	gtg Val 860	ccc Pro	gcc Ala	gac Asp	cgc Arg	2592
atc Ile 865	ctg Leu	gac Asp	gaa Glu	gcc Ala	aag Lys 870	aag Lys	cac His	aag Lys	gtc .Val	gac Asp 875	atg Met	atc Ile	ggc Gly	ctg Leu	tcg Ser 880	2640
ggc	ctg Leu	atc Ile	acc Thr	ccc Pro 885	tcg Ser	ctg Leu	gac Asp	gag Glu	atg Met 890	gtg Val	ttc Phe	gtg Val	gcc Ala	gcc Ala 895	gag Glu	2688
atg Met	gag Glu	cgc Arg	cag Gln 900	ggc Gly	ttt Phe	gat Asp	atc Ile	ccg Pro 905	ctg Leu	ctg Leu	atc Ile	ggc	ggc Gly 910	gcc Ala	acc Thr	2736
acc Thr	agc Ser	cgc Arg 915	acc Thr	cac His	acc Thr	gcg Ala	gtg Val 920	aag Lys	atc Ile	gag Glu	ccg Pro	gcc Ala 925	tat Tyr	cgc Arg	cgg Arg	2784
ggt Gly	ccg Pro 930	acg Thr	acc Thr	tat Tyr	gtc Val	gtc Val 935	gac Asp	gcc Ala	agc Ser	cgc Arg	gcc Ala 940	gtg Val	ggc Gly	gtg Val	gtc Val	2832
tcg Ser 945	ggc Gly	ctg Leu	ctg Leu	tcg Ser	gaa Glu 950	ggc Gly	gag Glu	cgt Arg	gac Asp	cgg Arg 955	atc Ile	atc Ile	gcc Ala	gag Glu	acc Thr 960	2880
cgc Arg	gcc Ala	gag Glu	tat Tyr	gtg Val 965	aag Lys	gtc Val	cgc Arg	gag Glu	caa Gln 970	tac Tyr	gcg Ala	cgc Arg	ggc Gly	cag Gln 975	acc Thr	2928
acc Thr	aag Lys	gcc Ala	cgc Arg 980	gcc Ala	tcg Ser	atc Ile	cag Gln	gag Glu 985	gcc Ala	cgc Arg	aag Lys	cgc Arg	gcc Ala 990	ttc Phe	gcc Ala	2976
att Ile	gac Asp	tgg Trp 995	aag Lys	ggc Gly	tat Tyr	Ala	ccg Pro	ccc Pro	aag Lys	ccc Pro	Ala	ttc Phe 1005	atc Ile	ggc	acg Thr	3024
Arg	gtg Val 010	ttc Phe	gag Glu	ccg Pro	tcg Ser 1	ctg Leu 015	gcc Ala	gag Glu	ctg Leu	Val	ccg Pro 1020	ttc Phe	atc Ile	gac Asp	tgg Trp	3072
tcg Ser 1025	Pro	ttc Phe	ttc Phe	gcc Ala	agc Ser 1030	Trp	gag Glu	ctg Leu	atc Ile	ggc Gly 1035	Arg	ttc Phe	ccg Pro	cag Gln	atc Ile 1040	3120
ctg Leu	gag Glu	gac Asp	qaA	gtg Val 045	gtc Val	ggc Gly	cag Gln	Ala	gcc Ala .050	acc Thr	gac Asp	ctc Leu	Tyr	cgc Arg 1055	gac Asp	3168
gcc Ala	cgc Arg	Ala	atg Met 060	ctg Leu	gac Asp	aag Lys	Val	gtc Val .065	gag Glu	gaa Glu	aag Lys	Trp	ttc Phe 1070	999 999	gcc Ala	3216
aag Lys	ggc Gly	gtg Val	atc Ile	ggc Gly	ttc Phe	tgg Trp	ccg Pro	gcc Ala	cag Gln	gcc Ala	cag Gln	ggc Gly	gac Asp	gac Asp	atc Ile	3264

175 1075 1080 1085 gtg ctc tat acc gac gag acc cgc gtg gcc gag ttc tcg cgc ctg cac 3312 Val Leu Tyr Thr Asp Glu Thr Arg Val Ala Glu Phe Ser Arg Leu His 1090 1095 acc ctt cgc cag cag atg gac aag ggc gcc gac aag agc ggc gag gcc Thr Leu Arg Gln Gln Met Asp Lys Gly Ala Asp Lys Ser Gly Glu Ala 1110 1115 aag gcc aat gtc gcc ctg tcg gac ttc gtc gcg ccg atc ggg cag ggg 3408 Lys Ala Asn Val Ala Leu Ser Asp Phe Val Ala Pro Ile Gly Gln Gly 1125 get gae tat gte gge gge tte gee gte ace gea gge cat gge gag gae Ala Asp Tyr Val Gly Gly Phe Ala Val Thr Ala Gly His Gly Glu Asp gag atc gtc gcc aag ttc aag gcg gcc ggc gac gac tac aac gcc atc 3504 Glu Ile Val Ala Lys Phe Lys Ala Ala Gly Asp Asp Tyr Asn Ala Ile 1160 1165 atg gcc tcg gcc ctg gcc gac cgc ctg gcc gaa gcc ttc gcc gag tgg 3552 Met Ala Ser Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Trp 1175 ctg cac tac aaa gcc cgt gtc gag ctg tgg ggc tac gcc gcc gac gag 3600 Leu His Tyr Lys Ala Arg Val Glu Leu Trp Gly Tyr Ala Ala Asp Glu 1190 1195 gac gcc gac gtc gag cgc ctg atc gcc gaa aag tac cag ggc atc cgc 3648 Asp Ala Asp Val Glu Arg Leu Ile Ala Glu Lys Tyr Gln Gly Ile Arg 1205 1210 ecc geg ecc gge tat eeg gee eag ecc gae eac acc gag aaa ggt acg 3696 Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Gly Thr 1220 1225 ctg ttc aag ctg ctc gac gcc gag gcg gcc acc ggt ctg cag ctg acc 3744 Leu Phe Lys Leu Leu Asp Ala Glu Ala Ala Thr Gly Leu Gln Leu Thr 1235 1240 gag age tac gee atg ace cet gge geg geg gte tee gge etg tte tte 3792 Glu Ser Tyr Ala Met Thr Pro Gly Ala Ala Val Ser Gly Leu Phe Phe 1250 1255 age car ege cag geg car tat the ggg ghe gge aag ate gar ger qar Ser His Arg Gln Ala His Tyr Phe Gly Val Gly Lys Ile Asp Ala Asp 1265 1270 cag gtc gag gac tac gcc cgc cgc aag ggc tgg gat atg gag acg qcc 3888 Gln Val Glu Asp Tyr Ala Arg Arg Lys Gly Trp Asp Met Glu Thr Ala 1285 gag cgc tgg ctg tcg ccg atc ctg aac tac gat ccg cta gcg cgg gcg 3936 Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Asp Pro Leu Ala Arg Ala 1300 1305 cgc ggg gcg gct tag 3954 Arg Gly Ala Ala Ala

<210> 40

<211> 1317

<212> PRT

<213> Caulobacter_crescentus

<220>

<221> unsure

<222> 359 .. 359

<223> All occurrences of Xaa indicate any amino acid ...

<220>

<221> unsure

<222> 416 .. 416

<223> All occurrences of Xaa indicate any amino acid

Met Thr Asp Leu Ser Ile Arg Ala Asn Arg Val Ala Ala Leu Lys Ala

Ala Ala Lys Glu Arg Ile Leu Ile Leu Asp Gly Ser Trp Gly Val Met 25 30

Phe Gln Lys Lys Gly Leu Thr Glu Ala Asp Tyr Arg Ala Glu Arg Phe 40

Ala Ala Tyr Asn Gly Gln Met Lys Gly Asn Asn Asp Ile Leu Cys Leu

Thr Arg Pro Asp Leu Val Ala Glu Leu His Asp Ala Tyr Phe Ser Ala

Gly Ala Asp Ile Ser Glu Thr Asn Thr Phe Ser Gly Thr Thr Ile Ala

Gln Ala Asp Tyr His Leu Gly Glu Gln Asp Val Trp Asp Ile Asn Leu 100 105

Glu Gly Ala Lys Ile Gly Arg Ser Val Ala Asp Arg Trp Asn Ala Gln

Asn Pro Asp Arg Pro Lys Phe Ile Ala Gly Ser Met Gly Pro Leu Asn 130 135

Val Met Leu Ser Met Ser Ser Asp Val Asn Asp Pro Gly Ala Arg Lys 150

Val Thr Phe Asp Gln Val Tyr Glu Ala Tyr Arg Gln Gln Val Asp Ala

Leu Tyr Gln Gly Gly Val Asp Leu Phe Leu Ile Glu Thr Ile Thr Asp

Thr Leu Asn Cys Lys Ala Ala Ile Lys Ala Ile Leu Asp Trp Arg Asp

Glu Gly His Glu Glu Leu Pro Ile Trp Ile Ser Gly Thr Ile Thr Asp 210 215

Arg Ser Gly Arg Thr Leu Ser Gly Gln Thr Ala Glu Ala Phe Trp Asn 225 230 235

177

Ser Val Lys His Ala Lys Pro Phe Ala Val Gly Phe Asn Cys Ala Leu 245 250 255

- Gly Ala Asp Leu Met Arg Pro His Ile Ala Glu Met Ala Arg Ile Ala 260 265 270
- Asp Thr Leu Val Ala Ala Tyr Pro Asn Ala Gly Leu Pro Asn Ala Met
 275 280 285
- Gly Gln Tyr Asp Glu Glu Pro His Glu Thr Gly His Ala Leu His Glu 290 295 300
- Trp Ala Lys Asp Gly Leu Val Asn Ile Leu Gly Gly Cys Cys Gly Thr 305 310 315 320
- Thr Pro Asp His Ile Arg His Val Ala Asp Glu Val Arg Gly Val Thr 325 330 335
- Pro Arg Gln Ile Pro Glu Arg Pro Lys Ala Met Arg Leu Ala Gly Leu 340 345 350
- Glu Pro Phe Glu Leu Ala Xaa Trp Leu Arg Pro Gln Ile Pro Phe Ser 355 360 365
- Pro Cys Gly Arg Arg Cys Arg Arg Arg Arg Arg Met Arg Gly Leu Ala 370 380
- Gly Pro Ser Thr Ala Val Ser Arg Arg Arg Ser Ser Thr Pro His 385 390 395 400
- Pro Thr Arg Cys Ala Gly His Leu Leu Pro Gln Gly Glu Lys Gly Xaa 405 410 415
- Leu Leu Leu Asp Pro Glu Met Arg Pro Val Phe Val Asn Ile Gly Glu 420 425 430
- Arg Thr Asn Val Thr Gly Ser Ala Lys Phe Lys Lys Leu Ile Val Glu 435 440 445
- Gly Asn Tyr Pro Glu Ala Leu Ser Val Ala Arg Gln Gln Val Glu Ala 450 460
- Gly Ala Gln Val Ile Asp Val Asn Met Asp Glu Gly Leu Leu Asp Ser 465 470 475 480
- Gln Gln Ala Met Val Thr Phe Leu Asn Leu Met Ala Ala Glu Pro Asp 485 490 495
- Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Trp Glu Val Ile 500 505 510
- Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile 515 520 525
- Ser Leu Lys Glu Gly Glu Glu Lys Phe Leu Glu Gln Ala Thr Leu Cys 530 540
- Leu Arg Tyr Gly Ala Ala Val Val Met Ala Phe Asp Glu Val Gly 555 560
- Gln Ala Asp Thr Glu Lys Arg Lys Val Glu Ile Cys Thr Arg Ala Tyr 565 570 575

- Asn Thr Leu Val Asp Lys Val Gly Phe Pro Pro Glu Asp Ile Ile Phe 580 585 590
- Asp Pro Asn Ile Phe Ala Val Ala Thr Gly Ile Glu Glu His Asp Asn 595 600 605
- Tyr Ala Val Asp Phe Ile Glu Ala Thr Arg Arg Ile Lys Gln Met Leu 610 615 620
- Pro Tyr Ala Arg Val Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe 625 630 635 640
- Arg Gly Asn Glu Pro Val Arg Arg Ala Ile His Ser Val Phe Leu Tyr
 645 650 655
- His Ala Ile Asn Ala Gly Met Asp Met Gly Ile Val Asn Ala Gly Asp
 660 665 670
- Leu Pro Val Tyr Asp Asp Ile Asp Pro Ala Leu Arg Glu Ala Val Glu 675 680 685
- Asp Val Ile Leu Asn Arg Pro Gln Arg Asp Pro Val Met Thr Asn Thr 690 695 700
- Glu Arg Leu Val Glu Met Ala Pro Arg Tyr Lys Gly Glu Lys Gly Gln 705 710 715 720
- Gln Gln Val Ala Asn Leu Glu Trp Arg Lys Gly Thr Val Asn Glu Arg 725 730 735
- Leu Thr His Ala Leu Val His Gly Ile Thr Glu Phe Ile Glu Gln Asp 740 745 750
- Thr Glu Glu Ala Arg Leu Ala Ala Glu Arg Pro Leu His Val Ile Glu 755 760 765
- Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala
 770 780
- Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys 785 790 795 800
- Gln Ala Val Ala Trp Leu Met Pro Phe Met Glu Ala Glu Lys Glu Gly 805 810 815
- Gln Glu Arg Lys Ala Ala Gly Lys Val Leu Met Ala Thr Val Lys Gly 820 825 830
- Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln Cys 835 840 845
- Asn Asn Tyr Glu Val Val Asp Leu Gly Val Met Val Pro Ala Asp Arg 850 855 860
- Ile Leu Asp Glu Ala Lys Lys His Lys Val Asp Met Ile Gly Leu Ser 865 870 875 880
- Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Phe Val Ala Ala Glu 885 890 895
- Met Glu Arg Gln Gly Phe Asp Ile Pro Leu Leu Ile Gly Gly Ala Thr

900

905

910

Thr Ser Arg Thr His Thr Ala Val Lys Ile Glu Pro Ala Tyr Arg Arg 915 920 925

Gly Pro Thr Thr Tyr Val Val Asp Ala Ser Arg Ala Val Gly Val Val 930 935 940

Ser Gly Leu Leu Ser Glu Gly Glu Arg Asp Arg Ile Ile Ala Glu Thr 945 950 955 960

Arg Ala Glu Tyr Val Lys Val Arg Glu Gln Tyr Ala Arg Gly Gln Thr 965 970 975

Thr Lys Ala Arg Ala Ser Ile Gln Glu Ala Arg Lys Arg Ala Phe Ala 980 985 990

Ile Asp Trp Lys Gly Tyr Ala Pro Pro Lys Pro Ala Phe Ile Gly Thr 995 1000 1005

Arg Val Phe Glu Pro Ser Leu Ala Glu Leu Val Pro Phe Ile Asp Trp 1010 1015 1020

Ser Pro Phe Phe Ala Ser Trp Glu Leu Ile Gly Arg Phe Pro Gln Ile 1025 1030 1035 1040

Leu Glu Asp Asp Val Val Gly Gln Ala Ala Thr Asp Leu Tyr Arg Asp 1045 1050 1055

Ala Arg Ala Met Leu Asp Lys Val Val Glu Glu Lys Trp Phe Gly Ala 1060 1065 1070

Lys Gly Val Ile Gly Phe Trp Pro Ala Gln Ala Gln Gly Asp Asp Ile 1075 1080 1085

Val Leu Tyr Thr Asp Glu Thr Arg Val Ala Glu Phe Ser Arg Leu His 1090 1095 1100

Thr Leu Arg Gln Gln Met Asp Lys Gly Ala Asp Lys Ser Gly Glu Ala 1105 1110 1115 1120

Lys Ala Asn Val Ala Leu Ser Asp Phe Val Ala Pro Ile Gly Gln Gly 1125 1130 1135

Ala Asp Tyr Val Gly Gly Phe Ala Val Thr Ala Gly His Gly Glu Asp 1140 1145 1150

Glu Ile Val Ala Lys Phe Lys Ala Ala Gly Asp Asp Tyr Asn Ala Ile 1155 1160 1165

Met Ala Ser Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Trp 1170 1175 1180

Leu His Tyr Lys Ala Arg Val Glu Leu Trp Gly Tyr Ala Ala Asp Glu 1185 1190 1195 1200

Asp Ala Asp Val Glu Arg Leu Ile Ala Glu Lys Tyr Gln Gly Ile Arg 1205 1210 1215

Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Gly Thr 1220 1225 1230

180

Leu Phe Lys Leu Leu Asp Ala Glu Ala Ala Thr Gly Leu Gln Leu Thr Glu Ser Tyr Ala Met Thr Pro Gly Ala Ala Val Ser Gly Leu. Phe Phe 1250 1255 Ser His Arg Gln Ala His Tyr Phe Gly Val Gly Lys Ile Asp Ala Asp 1270 1275 Gln Val Glu Asp Tyr Ala Arg Arg Lys Gly Trp Asp Met Glu Thr Ala 1285 1290 Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Asp Pro Leu Ala Arg Ala 1305 Arg Gly Ala Ala Ala 1315 <210> 41 <211> 3759 <212> DNA <213> Rhodobacter capsulatus <220> <221> CDS <222> (1) .. (3756) <223> RRC01731 <400> 41 atg ctg acc cag acc ctg ccc cga tct gcg gcc ttt gcc gca att gag Met Leu Thr Gln Thr Leu Pro Arg Ser Ala Ala Phe Ala Ala Ile Glu 1 5 gcg ctt tcg cgc cag cgg atc ttg atc ctt gac ggg gcg atg ggc acg 96 Ala Leu Ser Arg Gln Arg Ile Leu Ile Leu Asp Gly Ala Met Gly Thr 20 25 30 cag atc cag cag ctt ggc ctg agc gag gac gat ttt ctg ggc cac ggc 144 Gln Ile Gln Gln Leu Gly Leu Ser Glu Asp Asp Phe Leu Gly His Gly teg gge tge gee tge ege cat gee ace gat cat eeg caa aag gge aac 192 Ser Gly Cys Ala Cys Arg His Ala Thr Asp His Pro Gln Lys Gly Asn 50 55 aac gac ctg ctg gtg ctg acc cag ccg caa gcg atc gag gag atc cat Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ala Ile Glu Glu Ile His 65 70 ttc cgc tat gcg atg gcg ggg gcg gat atc gtc gag acg aac acc ttt 288 Phe Arg Tyr Ala Met Ala Gly Ala Asp Ile Val Glu Thr Asn Thr Phe teg gee ace ace ate geg cag gee gat tae ggg etg gaa age geg gtg 336 Ser Ala Thr Thr Ile Ala Gln Ala Asp Tyr Gly Leu Glu Ser Ala Val 100 105 ttc gac ctg aac gcc gcg ggg gcg cgg gtg gcg cgg gcg atg gac Phe Asp Leu Asn Ala Ala Gly Ala Arg Val Ala Arg Ala Ala Met Asp 115

										01						
cgc Arg	gcc Ala 130	gag Glu	gcc Ala	acc Thr	gac Asp	gga Gly 135	cgg Arg	cgc Arg	cgc Arg	ttc Phe	gtt Val 140	gcg Ala	999 999	gcg Ala	gtg Val	432
999 Gly 145	Pro	acg Thr	aac Asn	cgc Arg	acc Thr 150	gcc Ala	tcg Ser	ctc Leu	tcg Ser	ccc Pro 155	gat Asp	gtg Val	aac Asn	gac Asp	ccg Pro 160	480
ggc Gly	ttt Phe	cgc Arg	gcc Ala	gtc Val 165	acc Thr	ttc Phe	Asp	gat Asp	ctg Leu 170	cgc Arg	acg Thr	gcc Ala	tat Tyr	ggc Gly 175	cag Gln	528
cag Gln	gtg Val	cgc Arg	ggt Gly 180	ctg Leu	atc Ile	gcg Ala	GJY 999	ggc Gly 185	gcc Ala	gat Asp	atc Ile	ctg Leu	ctg Leu 190	atc Ile	gag Glu	576
acg Thr	atc Ile	ttt Phe 195	gac Asp	acg Thr	ctg Leu	aac Asn	gcc Ala 200	aag Lys	gcg Ala	gcg Ala	att Ile	ttc Phe 205	gcc Ala	tgt Cys.	ttc Phe	624
gaa Glu	gcc Ala 210	ttt Phe	gcc Ala	gaa Glu	cgg Arg	ggc Gly 215	gag Glu	cgg Arg	ctg Leu	ccg Pro	gtg Val 220	atg Met	att Ile	tcc Ser	ggc Gly	672
acg Thr 225	atc Ile	acc Thr	gat Asp	gcc Ala	tcg Ser 230	61Å 888	cgc Arg	aca Thr	ttg Leu	tcg Ser 235	99 9 Gly	cag Gln	acg Thr	ccg Pro	acc Thr 240	720
gcg Ala	ttc Phe	tgg Trp	cat His	tcg Ser 245	gtg Val	gct Ala	cat His	gcc Ala	cgg Arg 250	Pro	ttt Phe	acc Thr	gtg Val	999 Gly 255	ctg Leu	768
aac Asn	tgc Cys	gcg Ala	ctg Leu 260	ggc Gly	gcc Ala	agt Ser	gcg Ala	atg Met 265	cgt Arg	ccg Pro	cat His	ctg Leu	gcg Ala 270	gaa Glu	ctg Leu	816
gcg Ala	ggc	gtc Val 275	gcc Ala	ccc Pro	tgc Cys	gcg Ala	atc Ile 280	tgc Cys	gcc Ala	tat Tyr	ccc Pro	aat Asn 285	gcc Ala	G1y 999	ctg Leu	864
ccc Pro	aat Asn 290	gcc Ala	ttt Phe	ggc Gly	caa Gln	tat Tyr 295	gac Asp	gaa Glu	acc Thr	ccc Pro	gac Asp 300	cgg Arg	acc Thr	gcc Ala	gcg Ala	912
cag Gln 305	gtg Val	gcc Ala	gaa Glu	ttt Phe	gcc Ala 310	cgc Arg	gaa Glu	999 999	ctg Leu	gtc Val 315	aat Asn	gtc Val	gtg Val	ggc Gly	ggt Gly 320	960
tgc Cys	tgc Cys	ggc Gly	acc Thr	acc Thr 325	ccc Pro	gat Asp	cac His	atc Ile	cgc Arg 330	gcc Ala	atc Ile	gcg Ala	gaa Glu	gcc Ala 335	gtg Val	1008
aaa Lys	cct Pro	ttc Phe	ccg Pro 340	ccg Pro	agg Arg	gcc Ala	ctg Leu	cca Pro 345	agc Ser	cgt Arg	tat Tyr	ctg Leu	cgc Arg 350	ctt Leu	tcg Ser	1056
999 999	ctt Leu	gag Glu 355	ccc Pro	ttt Phe	acc Thr	ctg Leu	acg Thr 360	ccc Pro	gac Asp	att Ile	ccc Pro	ttc Phe 365	gtg Val	aac Asn	atc Ile	1104
ggc Gly	gag Glu	cgc Arg	acg Thr	aat Asn	gtc Val	acc Thr	ggc Gly	tcg Ser	gcc Ala	cgg Arg	ttc Phe	cgc Arg	aag Lys	atg Met	atc Ile	1152

	370					375			•	,	380					
		cgc Arg														1200
gaa Glu	aac Asn	ggc	gcg Ala	cag Gln 405	atc Ile	ctt Leu	gac Asp	atc Ile	aac Asn 410	atg Met	gac Asp	gag Glu	G1Å 333	ctg Leu 415	atc Ile	1248
		cag Gln														1296
ccc Pro	gac Asp	att Ile 435	gcc Ala	cgg Arg	gtg Val	ccg Pro	gtg Val 440	atg Met	atc Ile	gac Asp	agc Ser	tcg Ser 445	aaa Lys	tgg Trp	gag Glu	1344
		gag Glu														1392
_		agc Ser	_	_							_					1440
		ctg Leu														1488
		cag Gln														1536
		cgc Arg 515														1584
		gac Asp														1632
gac Asp 545	aat Asn	tac Tyr	ggc	Val	gat Asp 550	Phe	atc Ile	gag Glu	Ala	gct Ala 555	cgc	tgg Trp	atc Ile	cgg Arg	gcc Ala 560	1680
aac Asn	ctg Leu	ccg Pro	cat His	gcc Ala 565	cat His	gtc Val	tcg Ser	ggc Gly	999 Gly 570	gtg Val	tcg Ser	aac Asn	ctg Leu	tcc Ser 575	ttc Phe	1728
agc Ser	ttt Phe	cgc Arg	ggc Gly 580	aac Asn	gaa Glu	ccc Pro	gtg Val	cgc Arg 585	gcg Ala	gcg Ala	atg Met	cat His	gcg Ala 590	gtg Val	ttt Phe	1776
		cac His 595														1824
		ctg Leu			Tyr											1872

		,													_	Q 2, 22 00,
tgc	gag	gat	gtg	gtg	ctc	aac	cgc	cag	ccc	83 aaa	tcg	ggc	ggc	acc	gcg	1920
625					Leu 630					635					640	
Thr	Glu	Arg	Met	Leu 645	gag Glu	Val	Ala	Glu	Arg 650	Phe	Arg	Gly	Gly	Ala 655	Arg	1968
Glu	Glu	Lys	Thr 660	Arg	gat Asp	Leu	Ala	Trp 665	Arg	Asp	Trp	Pro	Val 670	Glu	Lys	2016
cgg Arg	ctc Leu	gaa Glu 675	cat His	gcg Ala	ctg Leu	gtc Val	aat Asn 680	ggc	atc Ile	acc Thr	gaa Glu	Phe 685	atc Ile	ga g Glu	gcc Ala	2064
					agg Arg											2112
Glu 705	Gly	Pro	Leu	Met	gcg Ala 710	Gly	Met	Asn	Val	Val 715	Gly	Asp	Leu	Phe	Gly 720	2160
gcg Ala	Gly	aag Lys	atg Met	ttc Phe 725	ctg Leu	cca Pro	cag Gln	gtg Val	gtg Val 730	aaa Lys	tcg Ser	gcg Ala	cgc Arg	gtg Val 735	atg Met	2208
Lys	Gln	Ala	Val 740	Ala	gtt Val	Leu	Leu	Pro 745	Tyr	Met	Asp	Ala	Glu 750	Lys	Ala	2256
Ala	Arg	Gly 755	Gly	Glu	GJ y 999	Arg	Glu 760	Thr	Ala	Gly	Lys	Ile 765	Leu	Met	Ala	2304
Thr	Val 770	Lys	Gly	Asp	gtg Val	Нів 775	Asp	Ile	Gly	Lys	Asn 780	Ile	Val	Gly	Val	2352
Val 785	Leu	Ala	Сув	Asn	aat Asn 790	Tyr	qaA	Ile	Val	Авр 795	Leu	Gly	Val	Met	Val 800	2400
Pro	Pro	Gln	Lys	Ile 805	ctg Leu	Glu	Val	Ala	Arg 810	Ala	Glu	Lys	Val	Asp 815	Ala	2448
atc Ile	999 999	ctt Leu	tcc Ser 820	999 Gly	ctg Leu	atc Ile	acg Thr	cca Pro 825	agc Ser	ctg Leu	gac Asp	gag Glu	atg Met 830	gtg Val	cat His	2496
ctg Leu	gcc Ala	gcg Ala 835	gaa Glu	atg Met	gag Glu	cgc Arg	gag Glu 840	ggc Gly	ttt Phe	gac Asp	att Ile	ccg Pro 845	ctg Leu	ctg Leu	atc Ile	2544
					tcg Ser											2592
gcc Ala	tac Tyr	agc Ser	cgc Arg	999 Gly	cag Gln	gcg Ala	gtt Val	tat Tyr	gtg Val	ctc Leu	gat Asp	gcc Ala	agc Ser	cgg Arg	gcc Ala	2640

0.55									1	84						
865					870					8,75					880	,
gtg Val	999 999	gtg Val	gtg Val	999 61y 885	gcg Ala	ctt Leu	ttg Leu	agc Ser	ccg Pro 890	Asn	cag Gln	aag Lys	gtc Val	gat Asp 895	tac Tyr	2688
gcg Ala	gcg Ala	cag Gln	atc Ile 900	cgc	gcg	gac Asp	tat Tyr	gcg Ala 905	cag Gln	atc Ile	gcc Ala	gcc Ala	cgt Arg 910	cat His	gcc Ala	2736
cgc Arg	gac Asp	gag Glu 915	gcc Ala	gcc Ala	aag Lys	gtg Val	cgg Arg 920	ctg Leu	cct	ttg Leu	gcc Ala	gcg Ala 925	gcc Ala	cgg Arg	gcc Ala	2784
aat Asn	gcg Ala 930	ctg Leu	cgg Arg	ctc Leu	gac Asp	tgg Trp 935	tcg Ser	ggc Gly	tat Tyr	gcc Ala	gtg Val 940	ccc Pro	gcg Ala	ccg Pro	caa Gln	2832
ttc Phe 945	ctt Leu	ggc	ccg Pro	cgc Arg	gtg Val 950	atc Ile	gac Asp	gac Asp	tgg Trp	gat Asp 955	ctg Leu	gcc Ala	gaa Glu	gtg Val	gcg Ala 960	2880
cgg Arg	tat Tyr	atc Ile	gac Asp	tgg Trp 965	acg Thr	ccc Pro	ttc Phe	ttc Phe	cat His 970	gcc Ala	tgg Trp	gaa Glu	ttg Leu	aag Lys 975	61Å 888	2928
gtc Val	tat Tyr	ccg Pro	cgg Arg 980	att Ile	ctc Leu	gat Asp	gac Asp	gcc Ala 985	gaa Glu	aag Lys	ggc Gly	gaa Glu	gcg Ala 990	Ala	cgg Arg	2976
gca Ala	ctt Leu	ttc Phe 995	gcc Ala	gat Asp	gcc Ala	Gln	gcg Ala 1000	atg Met	ctg Leu	gcg Ala	Gln	atc Ile 1005	att Ile	gcc Ala	gaa Glu	3024
Arg	tgg Trp 010	ttc Phe	acc Thr	ccg Pro	cgc Arg 1	gcc Ala 015	gtg Val	gtg Val	GJA aaa	Phe	tgg Trp 1020	ccc Pro	gcg Ala	cag Gln	gcg Ala	3072
gtg Val 1025	Gly	gac Asp	gat Asp	atc Ile	cgg Arg 1030	Leu	tac Tyr	acc Thr	gac Asp	gag Glu 1035	Ser	cgg Arg	acc Thr	gaa Glu	gac Asp 1040	3120
ctc Leu	gcc Ala	act Thr	Phe	ttc Phe 045	acc Thr	ctg Leu	cgc Arg	Gln	cag Gln 1050	acc Thr	ggc Gly	aag Lys	Arg	gaa Glu 055	ggc Gly	3168
cġc Arg	ccg Pro	Asn	gtg Val 060	gct Ala	ttg Leu	gcc Ala	Asp	ttc Phe 065	gtc Val	gcg Ala	cct Pro	Ala	ggc Gly 070	acg Thr	gtg Val	3216
ccc (Авр	tat Tyr 075	ctg Leu	ggc Gly	ggc Gly	Phe	gtg Val 080	gtc Val	acc Thr	gcg Ala	Gly	ccc Pro .085	gag Glu	gaa Glu	gcc Ala	3264
Glu :	atc Ile 090	gcc Ala	gcg Ala	cgg Arg	ttc Phe 1	gaa Glu 095	gct Ala	gcc Ala	aat Asn	Asp	cat His 100	tat Tyr	tcc Ser	gcg Ala	atc Ile	3312
ctg (Leu \ 1105	gtc Val	aag Lys	gcg (Ala :	ctg Leu	gcc Ala 1110	qaA	cgc Arg	ttt Phe	gcc Ala	gaa Glu 1115	Ala	ctg Leu	gcc Ala	gag Glu	gcc Ala 1120	3360

185 ctg cat cag cgg gtg cgg cgc gac tat tgg ggc tat gcg ccc gaa gaa	3408
Leu His Gln Arg Val Arg Arg Asp Tyr Trp Gly Tyr Ala Pro Glu Glu	
1125 1130 1135	
age tte gee eee gat cag etg gtg gge gag eee tat ege gge ate ege	3456
Ser Phe Ala Pro Asp Gln Leu Val Gly Glu Pro Tyr Arg Gly Ile Arg 1140 1145 1150	
	2504
ccg gcg ccc ggc tat ccg gcc cag ccc gac cac acg gaa aag ctg acg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Leu Thr	3504
1155 1160 1165	
ctg ttc cgg ctg ctt ggg gcc gag gcc gcg acc ggc gtg cat ctg acc	3552
Leu Phe Arg Leu Leu Gly Ala Glu Ala Ala Thr Gly Val His Leu Thr 1170 1175 1180	
gac agc atg gcg atg tgg ccc ggc tct tcg gtc tcg ggg ctc tat atc Asp Ser Met Ala Met Trp Pro Gly Ser Ser Val Ser Gly Leu Tyr Ile	3600
1185 1190 1195 1200	
ggc cat ccg gag gcc tat tat ttc ggt ctg gcc cgg atc gag cag gat	3648
Gly His Pro Glu Ala Tyr Tyr Phe Gly Leu Ala Arg Ile Glu Gln Asp	
1205 1210 1215	
cag gcc gcc gat tac gcc gcc cgc aag ggc atg gcc ttg gcc gag gtg Gln Ala Ala Asp Tyr Ala Ala Arg Lys Gly Met Ala Leu Ala Glu Val	3696
1220 1225 1230	
cag cgc tgg ctg gcc ccg gtg ctg ggg tcg gcc gcg ccc gcc g	3744
Gln Arg Trp Leu Ala Pro Val Leu Gly Ser Ala Ala Pro Ala Ala Ala	3,44
1235 1240 1245	
gcg gtg gcc gcg tga Ala Val Ala Ala	3759
1250	
<210> 42	
<211> 1252 <212> PRT	
<213> Rhodobacter capsulatus	
<400> 42	
Met Leu Thr Gln Thr Leu Pro Arg Ser Ala Ala Phe Ala Ala Ile Glu 1 5 10	
20	
Ala Leu Ser Arg Gln Arg Ile Leu Ile Leu Asp Gly Ala Met Gly Thr 20 25 30	
Gln Ile Gln Gln Leu Gly Leu Ser Glu Asp Asp Phe Leu Gly His Gly 35 40	
	
Ser Gly Cys Ala Cys Arg His Ala Thr Asp His Pro Gln Lys Gly Asn 50 55	
Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ala Ile Glu Glu Ile His 65 70 75 80	
Phe Arg Tyr Ala Met Ala Gly Ala Asp Ile Val Glu Thr Asn Thr Phe 85 90 95	
Ser Ala Thr Thr Ile Ala Gln Ala Asp Tyr Gly Leu Glu Ser Ala Val	
oce are the the sta off was usb Tal Gia per gir set yis sal	

100

105

110

Phe Asp Leu Asn Ala Ala Gly Ala Arg Val Ala Arg Ala Ala Met Asp 115 120 125

Arg Ala Glu Ala Thr Asp Gly Arg Arg Phe Val Ala Gly Ala Val 130 135 140

Gly Pro Thr Asn Arg Thr Ala Ser Leu Ser Pro Asp Val Asn Asp Pro 145 150 155 160

Gly Phe Arg Ala Val Thr Phe Asp Asp Leu Arg Thr Ala Tyr Gly Gln 165 170 175

Gln Val Arg Gly Leu Ile Ala Gly Gly Ala Asp Ile Leu Leu Ile Glu 180 185 190

Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Cys Phe 195 200 205

Glu Ala Phe Ala Glu Arg Gly Glu Arg Leu Pro Val Met Ile Ser Gly 210 215 220

Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Pro Thr 225 230 235 240

Ala Phe Trp His Ser Val Ala His Ala Arg Pro Phe Thr Val Gly Leu 245 250 255

Asn Cys Ala Leu Gly Ala Ser Ala Met Arg Pro His Leu Ala Glu Leu 260 265 270

Ala Gly Val Ala Pro Cys Ala Ile Cys Ala Tyr Pro Asn Ala Gly Leu 275 280 285

Pro Asn Ala Phe Gly Gln Tyr Asp Glu Thr Pro Asp Arg Thr Ala Ala 290 295 300

Gln Val Ala Glu Phe Ala Arg Glu Gly Leu Val Asn Val Val Gly Gly 305 310 315 320

Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Ile Ala Glu Ala Val 325 330 335

Lys Pro Phe Pro Pro Arg Ala Leu Pro Ser Arg Tyr Leu Arg Leu Ser 340 345 350

Gly Leu Glu Pro Phe Thr Leu Thr Pro Asp Ile Pro Phe Val Asn Ile 355 360 365

Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Arg Phe Arg Lys Met Ile 370 375 380

Val Ala Arg Asp Tyr Ala Ala Ala Leu Asp Val Ala Arg Asp Gln Val 385 390 395 400

Glu Asn Gly Ala Gln Ile Leu Asp Ile Asn Met Asp Glu Gly Leu Ile 405 410 415

Asp Ser Gln Ala Ala Met Val Ala Phe Leu Asn Leu Leu Ala Ala Glu 420 425 430

187

Pro Asp Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Trp Glu 435 440 445

Val Ile Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Pro Val Val Asn 450 455 460

Ser Ile Ser Leu Lys Glu Gly Glu Glu Ile Phe Arg His His Ala Ala 465 470 480

Leu Cys Leu Ala Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu
485 490 495

Glu Gly Gln Ala Asp Ser Phe Ala Arg Lys Thr Ser Ile Cys Ala Arg
500 505 510

Ala Tyr Arg Ile Leu Val Glu Glu Ile Gly Phe Pro Pro Glu Asp Ile 515 520 525

Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Glu Glu His 530 535 540

Asp Asn Tyr Gly Val Asp Phe Ile Glu Ala Ala Arg Trp Ile Arg Ala 545 550 555 560

Asn Leu Pro His Ala His Val Ser Gly Gly Val Ser Asn Leu Ser Phe 565 570 575

Ser Phe Arg Gly Asn Glu Pro Val Arg Ala Ala Met His Ala Val Phe 580 585 590

Leu Tyr His Ala Ile Arg Ala Gly Met Asp Met Gly Ile Val Asn Ala 595 600 605

Gly Gln Leu Val Val Tyr Asp Gln Ile Asp Pro Glu Leu Arg Gln Ala 610 615 620

Cys Glu Asp Val Val Leu Asn Arg Gln Pro Lys Ser Gly Gly Thr Ala 625 630 635 640

Thr Glu Arg Met Leu Glu Val Ala Glu Arg Phe Arg Gly Gly Ala Arg 645 650 655

Glu Glu Lys Thr Arg Asp Leu Ala Trp Arg Asp Trp Pro Val Glu Lys 660 665 670

Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Glu Phe Ile Glu Ala 675 680 685

Asp Thr Glu Ala Ala Arg Leu Leu Ala Glu Arg Pro Leu His Val Ile 690 695 700

Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp Leu Phe Gly 705 710 715 720

Ala Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met
725 730 735

Lys Gln Ala Val Ala Val Leu Leu Pro Tyr Met Asp Ala Glu Lys Ala 740 745 750

Ala Arg Gly Gly Glu Gly Arg Glu Thr Ala Gly Lys Ile Leu Met Ala 755 760 765

- Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val 770 780
- Val Leu Ala Cys Asn Asn Tyr Asp Ile Val Asp Leu Gly Val Met Val 785 790 795 800
- Pro Pro Gln Lys Ile Leu Glu Val Ala Arg Ala Glu Lys Val Asp Ala 805 810 815
- Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val His 820 825 830
- Leu Ala Ala Glu Met Glu Arg Glu Gly Phe Asp Ile Pro Leu Leu Ile 835 840 845
- Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Ile Ala Pro 850 855 860
- Ala Tyr Ser Arg Gly Gln Ala Val Tyr Val Leu Asp Ala Ser Arg Ala 865 870 880
- Val Gly Val Val Gly Ala Leu Leu Ser Pro Asn Gln Lys Val Asp Tyr 885 890 895
- Ala Ala Gln Ile Arg Ala Asp Tyr Ala Gln Ile Ala Ala Arg His Ala 900 905 910
- Arg Asp Glu Ala Ala Lys Val Arg Leu Pro Leu Ala Ala Ala Arg Ala 915 920 925
- Asn Ala Leu Arg Leu Asp Trp Ser Gly Tyr Ala Val Pro Ala Pro Gln 930 935 940
- Phe Leu Gly Pro Arg Val Ile Asp Asp Trp Asp Leu Ala Glu Val Ala 945 950 955 960
- Arg Tyr Ile Asp Trp Thr Pro Phe Phe His Ala Trp Glu Leu Lys Gly 965 970 975
- Val Tyr Pro Arg Ile Leu Asp Asp Ala Glu Lys Gly Glu Ala Ala Arg 980 985 990
- Ala Leu Phe Ala Asp Ala Gln Ala Met Leu Ala Gln Ile Ile Ala Glu 995 1000 1005
- Arg Trp Phe Thr Pro Arg Ala Val Val Gly Phe Trp Pro Ala Gln Ala 1010 1015 1020
- Val Gly Asp Asp Ile Arg Leu Tyr Thr Asp Glu Ser Arg Thr Glu Asp 1025 1030 1035 1040
- Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Thr Gly Lys Arg Glu Gly 1045 1050 1055
- Arg Pro Asn Val Ala Leu Ala Asp Phe Val Ala Pro Ala Gly Thr Val 1060 1065 1070
- Pro Asp Tyr Leu Gly Gly Phe Val Val Thr Ala Gly Pro Glu Glu Ala 1075 1080 1085
- Glu Ile Ala Ala Arg Phe Glu Ala Ala Asn Asp His Tyr Ser Ala Ile

189 1090 1095 1100 Leu Val Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala Leu Ala Glu Ala 1110 1115

Leu His Gln Arg Val Arg Arg Asp Tyr Trp Gly Tyr Ala Pro Glu Glu 1125 1130

Ser Phe Ala Pro Asp Gln Leu Val Gly Glu Pro Tyr Arg Gly Ile Arg 1145 1150

Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Leu Thr 1160 .

Leu Phe Arg Leu Leu Gly Ala Glu Ala Ala Thr Gly Val His Leu Thr 1170

Asp Ser Met Ala Met Trp Pro Gly Ser Ser Val Ser Gly Leu Tyr Ile 1190

Gly His Pro Glu Ala Tyr Tyr Phe Gly Leu Ala Arg Ile Glu Gln Asp

Gln Ala Ala Asp Tyr Ala Ala Arg Lys Gly Met Ala Leu Ala Glu Val 1225

Gln Arg Trp Leu Ala Pro Val Leu Gly Ser Ala Ala Pro Ala Ala Ala 1235 1240

Ala Val Ala Ala 1250

<210> 43

<211> 3798

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(3795)

<223> RHS24705

<400> 43

atg tca ccc gcg ctc caa gac ctg tcg caa ccc gaa ggt ctg aag aaa 48 Met Ser Pro Ala Leu Gln Asp Leu Ser Gln Pro Glu Gly Leu Lys Lys 10

acc ctg cgg gat gag atc aat gcc att ctg cag aag agg att atg gtg 96 Thr Leu Arg Asp Glu Ile Asn Ala Ile Leu Gln Lys Arg Ile Met Val

ctg gat gga ggg atg ggg acc atg atc cag cgg gag aag cta aac gaa Leu Asp Gly Gly Met Gly Thr Met Ile Gln Arg Glu Lys Leu Asn Glu

gaa cac ttc cga ggt cag gaa ttt aaa gat cat gcc agg ccg ctg aaa 192 Glu His Phe Arg Gly Gln Glu Phe Lys Asp His Ala Arg Pro Leu Lys

ggc aac aat gac att tta agt ata act cag cct gat gtc att tac caa Gly Asn Asn Asp Ile Leu Ser Ile Thr Gln Pro Asp Val Ile Tyr Gln

65					70				7	190 75					80	
atc Ile	cat His	aag Lys	gaa Glu	tac Tyr 85	ttg Leu	ctg Leu	gct Ala	G1y 999	gca Ala 90	Asp	atc Ile	att Ile	gaa Glu	aca Thr 95	aat Asn	288
act Thr	ttt Phe	agc Ser	agc Ser 100	Thr	agt Ser	att Ile	gcc Ala	caa Gln 105	gct Ala	gac Asp	tat Tyr	ggc	ctt Leu 110	gaa Glu	cac His	336
ttg Leu	gcc Ala	tac Tyr 115	cgg Arg	atg Met	aac Asn	atg Met	tgc Cys 120	tct Ser	gca Ala	gga Gly	gtg Val	gcc Ala 125	aga Arg	aaa Lys	gct Ala	384
gcc Ala	gag Glu 130	gag Glu	gta Val	act Thr	ctc Leu	cag Gln 135	aca Thr	gga Gly	att Ile	aag Lys	agg Arg 140	ttt Phe	gtg Val	gca Ala	999 Gly	432
gct Ala 145	ctg Leu	ggt Gly	ccg Pro	act Thr	aat Asn 150	aag Lys	aca Thr	ctc Leu	tct Ser	gtg Val 155	tcc Ser	cca Pro	tct Ser	gtg Val	gaa Glu 160	480
Arg	Pro	Asp	Tyr	Arg 165	aac Asn	Ile	Thr	Phe	Авр 170	Glu	Leu	Val	Glu	Ala 175	Tyr	528
Gln	Glu	Gln	Ala 180	Lys	gga Gly	Leu	Leu	Asp 185	Gly	Gly	Val	Asp	Ile 190	Leu	Leu	576
Ile	Glu	Thr 195	Ile	Phe	gat Asp	Thr	Ala 200	Asn	Ala	ГÀв	Ala	Ala 205	Leu	Phe	Ala	624
Leu	Gln 210	Asn	Leu	Phe	gag Glu	Glu 215	Lys	Tyr	Ala	Pro	Arg 220	Pro	Ile	Phe	Ile	672
Ser 225	Gly	Thr	Ile	Val	gat Asp 230	Lys	Ser	Gly	Arg	Thr 235	Leu	Ser	Gly	Gln	Thr 240	720
Gly	Glu	Gly	Phe	Val 245	atc Ile	Ser	Val	Ser	His 250	Gly	Glu	Pro	Leu	Сув 255	Ile	768
Gly	Leu	Asn	Сув 260	Ala	ttg Leu	Gly	Ala	Ala 265	Glu	Met	Arg	Pro	Phe 270	Ile	Glu	816
Ile	Ile	Gly 275	Lys	Сув	aca Thr	Thr	Ala 280	Tyr	Val	Leu	Сув	Tyr 285	Pro	Asn	Ala	864
Gly	Leu 290	Pro	Asn	Thr	ttt Phe	Gly 295	Asp	Tyr	qaA	Glu	Thr 300	Pro	Ser	Met	Met	912
gcc Ala 305	aag Lys	cac His	cta Leu	Lys	gat Asp 310	ttt Phe	gct Ala	atg Met	gat Asp	ggc Gly 315	ttg Leu	gtc Val	aat Asn	ata Ile	gtt Val 320	960

									1	91						
					tca Ser											1008
_				_	aag Lys		_	_			_		_		_	1056
					tct Ser											1104
			_		att Ile			_	-		_	_				1152
_		_			atc Ile 390	-	_				-	_	-	-	-	1200
Val	Ala	Lys	Val	Gln 405	gtg Val	Glu	Met	Gly	Ala 410	Gln	Val	Leu	Авр	Val 415	Asn	1248
Met	Asp	Asp	Gly 420	Met	cta Leu	Āsp	Gly	Pro 425	Ser	Ala	Met	Thr	Arg 430	Phe	Сув	1296
Asn	Leu	Ile 435	Ala	Ser	gag Glu	Pro	Авр 440	Ile	Ala	Lys	Val	Pro 445	Leu	Сув	Ile	1344
Asp	Ser 450	Ser	Asn	Phe		Val 455	Ile	Glu	Ala	Gly	Leu 460	Lys	Сув	Сув	Gln	1392
Gly 465	Lys	Сув	Ile	Val	aat Asn 470	Ser	Ile	Ser	Leu	Lys 475	Glu	Gly	Glu	Asp	Asp 480	1440
Phe	Leu	Glu	Lys	Ala 485	Arg	Lys	Ile	Lys	Lys 490	Tyr	ĞÎy	Ala	Ala	Met 495	•	1488
Val	Met	Ala	Phe 500	Авр	Glu	Glu	Gly	Gln 505	Ala	Thr	Glu	Thr	Asp 510	Thr		1536
Ile	Arg	Val 515	Сув	Thr	cgg Arg	Ala	Tyr 520	His	Leu	Leu	Val	Lys 525	Lys	Leu	Gly	1584
Phe	Asn 530	Pro	Asn	Asp	Ile	Ile 535	Phe	Asp	Pro	Asn	Ile 540	Leu	Thr	Ile	GJA 333	1632
Thr 545	Gly	Met	Glu	Glu	His 550	Asn	Leu	Tyr	Ala	Ile 555	Asn	Phe	Ile	His	gca Ala 560	1680
															ggt	1728

565	570		

				565					570					575		
ctt Leu	tcc Ser	aac Asn	ttg Leu 580	Ser	ttc Phe	tcc Ser	ttc Phe	cga Arg 585	gga Gly	atg Met	gaa Glu	gcc Ala	att Ile 590	cga .Arg	gaa Glu	1776
gca Ala	atg Met	cat His 595	GJA aaa	gtt Val	ttc Phe	ctt Leu	tac Tyr 600	cat His	gca Ala	atc Ile	aag Lys	tct Ser 605	ggc Gly	atg Met	gac Asp	1824
atg Met	999 Gly 610	ata Ile	gtg Val	aat Asn	gct Ala	gga Gly 615	aac Asn	ctc Leu	cct Pro	gtg Val	tat Tyr 620	gat Asp	gat Asp	atc Ile	cat His	1872
aag Lys 625	gaa Glu	ctt Leu	ctg Leu	cag Gln	ctc Leu 630	tgt Cys	gaa Glu	gat Asp	ctc Leu	atc Ile 635	tgg Trp	aat Asn	aaa Lys	gac Asp	cct Pro 640	1920
gag Glu	gcc Ala	act Thr	gag Glu	aag Lys 645	ctc Leu	tta Leu	cgt Arg	tat Tyr	gcc Ala 650	cag Gln	act Thr	caa Gln	ggc Gly	aca Thr 655	gga Gly	1968
Gly	Lys	Lys	Val 660	Ile	Gln	act Thr	Asp	Glu 665	Trp	Arg	Asn	Gly	Pro 670	Val	Glu	2016
gaa Glu	cgc Arg	ctt Leu 675	gag Glu	tat Tyr	gcc Ala	ctt Leu	gtg Val 680	aag Lys	ggc Gly	att Ile	gaa Glu	aaa Lys 685	cat His	att Ile	att Ile	2064
Glu	Asp 690	Thr	Glu	Glu	Ala	agg Arg 695	Leu	Asn	Gln	Lys	Lys 700	Tyr	Pro	Arg	Pro	2112
Leu 705	Asn	Ile	Ile	Glu	Gly 710	ccc Pro	Leu	Met	Asn	Gly 715	Met	Lys	Ile	Val	Gly 720	2160
gat Asp	ctt Leu	ttt Phe	gga Gly	gct Ala 725	gga Gly	aaa Lys	atg Met	ttt Phe	cta Leu 730	cct Pro	cag Gln	gtt Val	ata Ile	aag Lys 735	tca Ser	2208
gcc Ala	cgg Arg	gtt Val	atg Met 740	aag Lys	aag Lys	gct Ala	gtt Val	ggc Gly 745	His	ctt Leu	atc Ile	cct Pro	ttc Phe 750	atg Met	gaa Glu	2256
aaa Lys	gaa Glu	aga Arg 755	gaa Glu	gaa Glu	acc Thr	aga Arg	gtg Val 760	ctt Leu	aac Asn	ggc Gly	aca Thr	gta Val 765	gaa Glu	gaa Glu	gag Glu	2304
gac Asp	cct Pro 770	tac Tyr	cag Gln	ggc Gly	acc Thr	atc Ile 775	gtg Val	ctg Leu	gcc Ala	act Thr	gtt Val 780	aaa Lys	ggc Gly	gac Asp	gtg Val	2352
cac His 785	gac Asp	ata Ile	ggc	aag Lys	aac Asn 790	ata Ile	gtt Val	gga Gly	gta Val	gtc Val 795	ctt Leu	ggc	tgc Cys	aat Asn	aat Asn 800	2400
ttc Phe	cga Arg	gtt Val	Ile	gat Asp 805	tta Leu	gga Gly	gtc Val	atg Met	act Thr 810	cca Pro	tgt Cys	gat Asp	aag Lys	ata Ile 815	ctg Leu	2448

										93						
aaa Lys	gct	gct Ala	ctt Leu 820	gac Asp	cac His	aaa Lys	gca Ala	gat Asp 825	ata Ile	att Ile	ggc Gly	ctg Leu	tca Ser 830	gga Glý	ctc Leu	2496
atc Ile	act Thr	cct Pro 835	tcc Ser	ctg Leu	gat Asp	gaa Glu	atg Met 840	att Ile	ttt Phe	gtt Val	gcc Ala	aag Lys 845	gaa Glu	atg Met	gag Glu	2544
aga Arg	tta Leu 850	gct Ala	ata Ile	agg Arg	att Ile	cca Pro 855	ttg Leu	ttg Leu	att Ile	gga Gly	gga Gly 860	gca Ala	acc Thr	act Thr	tca Ser	2592
ааа L ув 865	acc Thr	cac His	aca Thr	gca Ala	gtt Val 870	aaa Lys	ata Ile	gct Ala	ccg Pro	aga Arg 875	tac Tyr	agt Ser	gca Ala	cct Pro	gta Val 880	2640
atc Ile	cat His	gtc Val	ctg Leu	gac Asp 885	gcg Ala	tcc Ser	aag Lys	agt Ser	gtg Val 890	gtg Val	gtg Val	tgt Cys	tcc Ser	cag Gln 895	ctg Leu	2688
tta Leu	gat Asp	gaa Glu	aat Asn 900	cta Leu	aag Lys	gat Asp	gaa Glu	tac Tyr 905	ttt Phe	gag Glu	gaa Glu	atc Ile	atg Met 910	gaa Glu	gaa Glu	2736
tat Tyr	gaa Glu	gat Asp 915	att Ile	aga Arg	cag Gln	gac Asp	cat His 920	tat Tyr	gag Glu	tct Ser	ctc Leu	aag Lys 925	gag Glu	agg Arg	aga Arg	2784
tac Tyr	tta Leu 930	ccc Pro	tta Leu	agt Ser	caa Gln	gcc Ala 935	aga Arg	aaa Lys	agt Ser	ggt Gly	ttc Phe 940	caa Gln	atg Met	gat Asp	tgg Trp	2832
ctg Leu 945	tct Ser	gaa Glu	cct Pro	cac His	cca Pro 950	gtg Val	aag Lys	ccc Pro	acg Thr	ttt Phe 955	att Ile	ggg Gly	acc Thr	cag Gln	gtc Val 960	2880
ttt Phe	gaa Glu	gac Asp	tat Tyr	gac Asp 965	ctg Leu	cag Gln	aag Lys	ctg Leu	gtg Val 970	gac Asp	tac Tyr	att Ile) Asp	tgg Trp 975	aag Lys	2928
cct Pro	ttc Phe	ttt Phe	gat Asp 980	gtc Val	tgg Trp	cag Gln	ctc Leu	cgg Arg 985	ggc Gly	aag Lys	tac Tyr	ccg Pro	aat Asn 990	cga Arg	ggc Gly	2976
ttt Phe	ccc Pro	aag Lys 995	ata Ile	ttt Phe	aac Asn	qaA	aaa Lys 000	aca Thr	gta Val	ggt Gly	Gly	gag Glu .005	gcc Ala	agg Arg	aag Lys	3024
Val	tac Tyr 010	gat Asp	gat Asp	gcc Ala	His	aat Asn 015	atg Met	ctg Leu	aac Asn	Thr	ctg Leu .020	att Ile	agt Ser	caa Gln	aag Lys	3072
aaa Lys 1025	Leu	cgg Arg	gcc Ala	cgg Arg	ggt Gly 1030	Val	gtt Val	Gly 999	ttc Phe	tgg Trp 1035	Pro	gca Ala	cag Gln	agt Ser	atc Ile 1040	3120
caa Gln) Asp	gac Asp	Ile	cac His 045	ctg Leu	tac Tyr	gcg Ala	Glu	gct Ala 050	gct Ala	gtg Val	ccc Pro	Gln	gct Ala 055	gca Ala	3168
gag Glu	ccc Pro	ata Ile	gcc Ala	acc Thr	ttc Phe	tat Tyr	ggg Gly	tta Leu	agg Arg	caa Gln	cag Gln	gct Ala	gag Glu	aag Lys	gac Asp	3216

1060 1065 1070 tot god age acg gag coa tac tac tgc ctc tca gac ttc atc got ccc Ser Ala Ser Thr Glu Pro Tyr Tyr Cys Leu Ser Asp Phe Ile Ala Pro ttg cat tct ggc atc cgt gac tac ctg ggc ctg ttt gcc gtt gcc tgc 3312 Leu His Ser Gly Ile Arg Asp Tyr Leu Gly Leu Phe Ala Val Ala Cys 1095 ttt ggg gta gaa gag ctg agc aag gcc tat gag gat gat ggt gac gac Phe Gly Val Glu Glu Leu Ser Lys Ala Tyr Glu Asp Asp Gly Asp Asp 1115 tac age age ate atg gte aag geg etg ggg gae egg etg gea gag gee 3408 Tyr Ser Ser Ile Met Val Lys Ala Leu Gly Asp Arg Leu Ala Glu Ala 1130 ttt gca gaa gag ctc cat gaa aga gtt cgc cga gaa ctg tgg gcc tac 3456 Phe Ala Glu Glu Leu His Glu Arg Val Arg Arg Glu Leu Trp Ala Tyr 1140 1145 tgt ggc agt gag cag ctg gac gtc gca gac ctg cgc agg ctg cgg tac 3504 Cys Gly Ser Glu Gln Leu Asp Val Ala Asp Leu Arg Arg Leu Arg Tyr 1155 1160 1165 aag ggc atc cgc ccg gct cct ggc tac ccc agc cag ccc gac cac acc 3552 Lys Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ser Gln Pro Asp His Thr 1170 1175 gag aag ctc acc atg tgg aga ctt gca gac atc gag cag tct aca ggc 3600 Glu Lys Leu Thr Met Trp Arg Leu Ala Asp Ile Glu Gln Ser Thr Gly 1185 1190 1195 att agg tta aca gaa tca tta gca atg gca cct gct tca gca gtc tca 3648 Ile Arg Leu Thr Glu Ser Leu Ala Met Ala Pro Ala Ser Ala Val Ser 1205 1210 ggc ctc tac ttc tcc aat ttg aag tcc aaa tat ttt gct gtg ggg aag 3696 Gly Leu Tyr Phe Ser Asn Leu Lys Ser Lys Tyr Phe Ala Val Gly Lys 1220 1225 att tcc aag gat cag gtt gag gat tat gca ttg agg aag aac ata tct 3744 Ile Ser Lys Asp Gln Val Glu Asp Tyr Ala Leu Arg Lys Asn Ile Ser 1235 gtg gct gag gtt gag aaa tgg ctt gga ccc att ttg gga tat gat aca 3792 Val Ala Glu Val Glu Lys Trp Leu Gly Pro Ile Leu Gly Tyr Asp Thr 1250 1255 gac taa 3798 Asp 1265 <210> 44 <211> 1265 <212> PRT <213> Homo sapiens

<400> 44

Met Ser Pro Ala Leu Gln Asp Leu Ser Gln Pro Glu Gly Leu Lys Lys

195 10 Thr Leu Arg Asp Glu Ile Asn Ala Ile Leu Gln Lys Arg Ile Met Val 25 Leu Asp Gly Gly Met Gly Thr Met Ile Gln Arg Glu Lys Leu Asn Glu Glu His Phe Arg Gly Gln Glu Phe Lys Asp His Ala Arg Pro Leu Lys Gly Asn Asn Asp Ile Leu Ser Ile Thr Gln Pro Asp Val Ile Tyr Gln 70 Ile His Lys Glu Tyr Leu Leu Ala Gly Ala Asp Ile Ile Glu Thr Asn 90 Thr Phe Ser Ser Thr Ser Ile Ala Gln Ala Asp Tyr Gly Leu Glu His 105 Leu Ala Tyr Arg Met Asn Met Cys Ser Ala Gly Val Ala Arg Lys Ala Ala Glu Glu Val Thr Leu Gln Thr Gly Ile Lys Arg Phe Val Ala Gly 135 Ala Leu Gly Pro Thr Asn Lys Thr Leu Ser Val Ser Pro Ser Val Glu 150 Arg Pro Asp Tyr Arg Asn Ile Thr Phe Asp Glu Leu Val Glu Ala Tyr 170 Gln Glu Gln Ala Lys Gly Leu Leu Asp Gly Gly Val Asp Ile Leu Leu Ile Glu Thr Ile Phe Asp Thr Ala Asn Ala Lys Ala Ala Leu Phe Ala Leu Gln Asn Leu Phe Glu Glu Lys Tyr Ala Pro Arg Pro Ile Phe Ile Ser Gly Thr Ile Val Asp Lys Ser Gly Arg Thr Leu Ser Gly Gln Thr 230 235 Gly Glu Gly Phe Val Ile Ser Val Ser His Gly Glu Pro Leu Cys Ile 250 Gly Leu Asn Cys Ala Leu Gly Ala Ala Glu Met Arg Pro Phe Ile Glu Ile Ile Gly Lys Cys Thr Thr Ala Tyr Val Leu Cys Tyr Pro Asn Ala 280 Gly Leu Pro Asn Thr Phe Gly Asp Tyr Asp Glu Thr Pro Ser Met Met Ala Lys His Leu Lys Asp Phe Ala Met Asp Gly Leu Val Asn Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Asp His Ile Arg Glu Ile Ala Glu 325

196

Ala Val Lys Asn Cys Lys Pro Arg Val Pro Pro Ala Thr Ala Phe Glu 340 345 350

Gly His Met Leu Leu Ser Gly Leu Glu Pro Phe Arg Ile Gly Pro Tyr 355 360 365

Thr Asn Phe Val Asn Ile Gly Glu Arg Cys Asn Val Ala Gly Ser Arg 370 375 380

Lys Phe Ala Lys Leu Ile Met Ala Gly Asn Tyr Glu Glu Ala Leu Cys 385 390 395 400

Val Ala Lys Val Gln Val Glu Met Gly Ala Gln Val Leu Asp Val Asn 405 410 415

Met Asp Asp Gly Met Leu Asp Gly Pro Ser Ala Met Thr Arg Phe Cys 420 425 430

Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Lys Val Pro Leu Cys Ile 435 440 445

Asp Ser Ser Asn Phe Ala Val Ile Glu Ala Gly Leu Lys Cys Cys Gln 450 460

Gly Lys Cys Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Asp Asp 465 470 475 480

Phe Leu Glu Lys Ala Arg Lys Ile Lys Lys Tyr Gly Ala Ala Met Val 485 490 495

Val Met Ala Phe Asp Glu Glu Gly Gln Ala Thr Glu Thr Asp Thr Lys
500 505 510

Ile Arg Val Cys Thr Arg Ala Tyr His Leu Leu Val Lys Lys Leu Gly
515 520 525

Phe Asn Pro Asn Asp Ile Ile Phe Asp Pro Asn Ile Leu Thr Ile Gly 530 540

Thr Gly Met Glu Glu His Asn Leu Tyr Ala Ile Asn Phe Ile His Ala 545 550 555 560

Thr Lys Val Ile Lys Glu Thr Leu Pro Gly Ala Arg Ile Ser Gly Gly 565 570 575

Leu Ser Asn Leu Ser Phe Ser Phe Arg Gly Met Glu Ala Ile Arg Glu 580 585 590

Ala Met His Gly Val Phe Leu Tyr His Ala Ile Lys Ser Gly Met Asp 595 600 605

Met Gly Ile Val Asn Ala Gly Asn Leu Pro Val Tyr Asp Asp Ile His 610 620

Lys Glu Leu Leu Gln Leu Cys Glu Asp Leu Ile Trp Asn Lys Asp Pro 625 630 640

Glu Ala Thr Glu Lys Leu Leu Arg Tyr Ala Gln Thr Gln Gly Thr Gly
645 650

Gly Lys Lys Val Ile Gln Thr Asp Glu Trp Arg Asn Gly Pro Val Glu 660 665 670

- Glu Arg Leu Glu Tyr Ala Leu Val Lys Gly Ile Glu Lys His Ile Ile 675 680 685
- Glu Asp Thr Glu Glu Ala Arg Leu Asn Gln Lys Lys Tyr Pro Arg Pro 690 695 700
- Leu Asn Ile Ile Glu Gly Pro Leu Met Asn Gly Met Lys Ile Val Gly 705 710 715 720
- Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Ile Lys Ser
 725 730 735
- Ala Arg Val Met Lys Lys Ala Val Gly His Leu Ile Pro Phe Met Glu
 740 745 750
- Lys Glu Arg Glu Glu Thr Arg Val Leu Asn Gly Thr Val Glu Glu Glu 755 760 765
- Asp Pro Tyr Gln Gly Thr Ile Val Leu Ala Thr Val Lys Gly Asp Val 770 780
- His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gly Cys Asn Asn 785 790 795 800
- Phe Arg Val Ile Asp Leu Gly Val Met Thr Pro Cys Asp Lys Ile Leu 805 810 815
- Lys Ala Ala Leu Asp His Lys Ala Asp Ile Ile Gly Leu Ser Gly Leu 820 825 830
- Ile Thr Pro Ser Leu Asp Glu Met Ile Phe Val Ala Lys Glu Met Glu 835 840 845
- Arg Leu Ala Ile Arg Ile Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser
- Lys Thr His Thr Ala Val Lys Ile Ala Pro Arg Tyr Ser Ala Pro Val 865 870 875 880
- Ile His Val Leu Asp Ala Ser Lys Ser Val Val Val Cys Ser Gln Leu 885 890 895
- Leu Asp Glu Asn Leu Lys Asp Glu Tyr Phe Glu Glu Ile Met Glu Glu 900 905 910
- Tyr Glu Asp Ile Arg Gln Asp His Tyr Glu Ser Leu Lys Glu Arg Arg 915 920 925
- Tyr Leu Pro Leu Ser Gln Ala Arg Lys Ser Gly Phe Gln Met Asp Trp 930 935 940
- Leu Ser Glu Pro His Pro Val Lys Pro Thr Phe Ile Gly Thr Gln Val 945 950 955 960
- Phe Glu Asp Tyr Asp Leu Gln Lys Leu Val Asp Tyr Ile Asp Trp Lys
 965 970 975
- Pro Phe Phe Asp Val Trp Gln Leu Arg Gly Lys Tyr Pro Asn Arg Gly 980 985 990
- Phe Pro Lys Ile Phe Asn Asp Lys Thr Val Gly Gly Glu Ala Arg Lys

995

1000

1005

Val Tyr Asp Asp Ala His Asn Met Leu Asn Thr Leu Ile Ser Gln Lys 1010 1015 1020

Lys Leu Arg Ala Arg Gly Val Val Gly Phe Trp Pro Ala Gln Ser Ile 1025 1030 1035 1040

Gln Asp Asp Ile His Leu Tyr Ala Glu Ala Ala Val Pro Gln Ala Ala 1045 1050 - 1055

Glu Pro Ile Ala Thr Phe Tyr Gly Leu Arg Gln Gln Ala Glu Lys Asp 1060 1065 1070

Ser Ala Ser Thr Glu Pro Tyr Tyr Cys Leu Ser Asp Phe Ile Ala Pro 1075 1080 1085

Leu His Ser Gly Ile Arg Asp Tyr Leu Gly Leu Phe Ala Val Ala Cys 1090 1095 1100

Phe Gly Val Glu Glu Leu Ser Lys Ala Tyr Glu Asp Asp Gly Asp Asp 1105 1110 1115 1120

Tyr Ser Ser Ile Met Val Lys Ala Leu Gly Asp Arg Leu Ala Glu Ala 1125 1130 1135

Phe Ala Glu Glu Leu His Glu Arg Val Arg Arg Glu Leu Trp Ala Tyr 1140 1145 1150

Cys Gly Ser Glu Gln Leu Asp Val Ala Asp Leu Arg Arg Leu Arg Tyr 1155 1160 1165

Lys Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ser Gln Pro Asp His Thr 1170 1175 1180

Glu Lys Leu Thr Met Trp Arg Leu Ala Asp Ile Glu Gln Ser Thr Gly
1185 1190 1195 1200

Ile Arg Leu Thr Glu Ser Leu Ala Met Ala Pro Ala Ser Ala Val Ser 1205 1210 1215

Gly Leu Tyr Phe Ser Asn Leu Lys Ser Lys Tyr Phe Ala Val Gly Lys 1220 1225 1230

Ile Ser Lys Asp Gln Val Glu Asp Tyr Ala Leu Arg Lys Asn Ile Ser 1235 1240 1245

Val Ala Glu Val Glu Lys Trp Leu Gly Pro Ile Leu Gly Tyr Asp Thr 1250 1255 1260

Asp 1265

<210> 45

<211> 3681

<212> DNA

<213> Vibrio fisheri

<220>

<221> CDS

<222> (1)..(3678)

<223> AB039955

<40	0 > 4	5														
gtg	gca Ala	gga	agc Ser	aat Asn 5	Ile	aaa Lys	gta Val	caa Gln	ata Ile 10	gaa Glu	aag Lys	caa Gln	ctt Leu	tca Ser 15	gag Glu	48
cga Arg	att Ile	tta Leu	ttg Leu 20	att Ile	gat Asp	ggt Gly	ggt Gly	atg Met 25	ggc	acc Thr	atg Met	att Ile	caa Gln 30	ggt Gly	tat Tyr	96
aag Lys	ttt Phe	gaa Glu 35	gag Glu	aaa Lys	gat Asp	tat Tyr	aga Arg 40	ggg Gly	gga Gly	cgc	ttt Phe	aat Asn 45	caa Gln	tgg Trp	cat His	144
tgt Cys	gat Asp 50	ctt Leu	aaa Lys	ggt Gly	aac Asn	aat Asn 55	gat Asp	tta Leu	tta Leu	gtt Val	ctt Leu 60	tca Ser	caa Gln	cca Pro	caa Gln	192
att Ile 65	ata Ile	aga Arg	gat Asp	ata Ile	cac His 70	gaa Glu	gcc Ala	tat Tyr	tta Leu	gaa Glu 75	gct Ala	ggt Gly	gct Ala	gat Asp	atc Ile 80	240
Leu	Glu	Thr	aat Asn	Thr 85	Phe	Asn	Ala	Thr	Thr 90	Ile	Ala	Met	Ala	Авр 95	Tyr	288
Asp	Met	Glu	agc Ser 100	Leu	Ser	Glu	Glu	11e 105	Asn	Phe	Glu	Ala	Ala 110	Lys	Leu	336
Ala	Arg	Glu 115	gtt Val	Ala	Asp	Lys	Trp 120	Thr	Glu	Lys	Thr	Pro 125	Asn	Lys	Pro	384
Arg	Tyr 130	Val	gca Ala	Gly	Val	Leu 135	Gly	Pro	Thr	Asn	Arg 140	Thr	Сув	Ser	Ile	432
Ser 145	Pro	Asp	gta Val	Asn	Asp 150	Pro	Gly	Phe	Arg	Asn 155	Val	Ser	Phe	Āsp	Glu 160	480
tta Leu	gtc Val	gaa Glu	gct Ala	tat Tyr 165	tca Ser	gag Glu	tca Ser	act Thr	cga Arg 170	gca Ala	ctt Leu	att Ile	aga Arg	ggt Gly 175	ggt Gly	528
tca Ser	gat Asp	ctt Leu	atc Ile 180	ctc Leu	atc Ile	gaa Glu	act Thr	ata Ile 185	ttt Phe	gat Asp	aca Thr	tta Leu	aat Asn 190	gct Ala	aaa Lys	576
gcg Ala	tgt Сув	tct Ser 195	ttt Phe	gct Ala	gtt Val	gaa Glu	tct Ser 200	gtt Val	ttt Phe	gaa Glu	gag Glu	ctt Leu 205	ggt Gly	att Ile	act Thr	624
Leu	cct Pro 210	gtt Val	atg Met	att Ile	tca Ser	999 Gly 215	acc Thr	att Ile	acc Thr	gat Asp	gca Ala 220	tca Ser	gga Gly	aga Arg	aca Thr	672
tta Leu 225	tcg Ser	999 Gly	caa Gln	aca Thr	aca Thr 230	gaa Glu	gct Ala	ttt Phe	tat Tyr	aat Asn 235	gca Ala	tta Leu	aga Arg	cat His	gta Val 240	720

aaa Lys	cct Pro	att Ile	tct Ser	Phe 245	ggt	ctt Leu	aac Asn	tgt Cys	gca Ala 250	Leu	ggt	cct Pro	Авр	gaa Glu . 255	Leu	768
cgt Arg	gaa Glu	tat Tyr	gta Val 260	Ser	gag Glu	ctt Leu	tca Ser	cgt Arg 265	Ile	tct Ser	gaa Glu	tgt Cys	tat Tyr 270	gtt Val	tct Ser	816
gcg Ala	cac His	cca Pro 275	Asn	gct Ala	ggt Gly	ttg Leu	cct Pro 280	aat Asn	gca Ala	ttt Phe	ggt Gly	gag Glu 285	tat Tyr	gat Asp	tta Leu	864
tct Ser	ccc Pro 290	gaa Glu	gat Asp	atg Met	gct Ala	gag Glu 295	cat His	gtt Val	gcg Ala	gaa Glu	tgg Trp 300	gca Ala	agc Ser	agc Ser	gga Gly	912
ttt Phe 305	tta Leu	aat Asn	ctt Leu	att Ile	ggt Gly 310	999 999	tgt Cys	tgt Cys	ggc Gly	acc Thr 315	act Thr	cct Pro	gaa Glu	cat His	att Ile 320	960
cgt Arg	caa Gln	atg Met	gct Ala	tta Leu 325	gtt Val	gtt Val	gaa Glu	ggt Gly	gtg Val 330	aaa Lys	cct Pro	cga Arg	caa Gln	tta Leu 335	cct Pro	1008
gaa Glu	tta Leu	ccc Pro	gta Val 340	gct Ala	tgt Cys	cgt Arg	ctt Leu	tcc Ser 345	gga Gly	tta Leu	gag Glu	cct Pro	tta Leu 350	aca	ata Ile	1056
Glu	Lys	Asp 355	Ser	Leu	ttt Phe	Ile	Asn 360	Val	Gly	Glu	Arg	Thr 365	Asn	Val	Thr	1104
gga Gly	tct Ser 370	gca Ala	cgt Arg	ttt Phe	aaa Lys	cgc Arg 375	tta Leu	att Ile	aaa Lys	gaa Glu	gag Glu 380	ctt Leu	tat Tyr	gac Asp	gaa Glu	1152
gca Ala 385	cta Leu	agt Ser	gtt Val	gct Ala	caa Gln 390	gag Glu	caa Gln	gtt Val	gaa Glu	aac Asn 395	ggt Gly	gct Ala	caa Gln	att Ile	atc Ile 400	1200
gat Asp	atc Ile	aac Asn	atg Met	gat Asp 405	gaa Glu	ggc Gly	atg Met	ctt Leu	gat Asp 410	gct Ala	gaa Glu	gca Ala	tgt Cys	atg Met 415	gtt Val	1248
cgt Arg	ttt Phe	tta Leu	aat Asn 420	ctt Leu	tgt Cys	gca Ala	tca Ser	gaa Glu 425	cct Pro	gaa Glu	ata Ile	tct Ser	ааа L ув 430	gta Val	cca Pro	1296
gtg Val	atg Met	gtt Val 435	gat Asp	tct Ser	tct Ser	Lys	tgg Trp 440	gaa Glu	gta Val	att Ile	gaa Glu	gct Ala 445	gga Gly	tta Leu	aag Lys	1344
tgt Cys	att Ile 450	caa Gln	ggt Gly	aag Lys	01Å 333	ata Ile 455	gtt Val	aat Asn	tca Ser	atc Ile	tct Ser 460	tta Leu	aag Lys	gaa Glu	ggc Gly	1392
ааа L ув 465	gaa Glu	aag Lys	ttt Phe	Val	cat His 470	caa Gln	gcc Ala	aag Lys	Leu	ata Ile 475	cgt Arg	cgt Arg	tat Tyr	ggt Gly	gct Ala 480	1440
gca	gtg	atc	gtt	atg	gct	ttt	gat	gaa	gtt	ggc	caa	gcg	gac	act	cgg	1488

									2	201						
Ala	. Val	Ile	e Val	. Met 485	Ala	Phe	qaA s	Glu	Val 490		Gln	Ala	Asp	Thr 495	Arg	
gag Glu	cgt Arg	aaa Lys	att Ile 500	gaa Glu	att Ile	tgt Cys	acc Thr	aat Asn 505	Ala	tac Tyr	aat Asn	att Ile	tta Leu 510	Val	gat Asp	1536
gaa Glu	gtt Val	990 Gly 515	Phe	cca Pro	Pro	gaa Glu	gat Asp 520	Ile	att	ttt Phe	gac Asp	cct Pro 525	aat Asn	att Ile	ttt Phe	1584
gcg Ala	gtt Val 530	Ala	aca Thr	ggt	atc Ile	gat Asp 535	Glu	cat His	aat Asn	aac Asn	tat Tyr 540	gca Ala	gta Val	gac Asp	ttt Phe	1632
att Ile 545	gaa Glu	gcc Ala	gtt Val	ggt Gly	gat Asp 550	Ile	aag Lys	cga Arg	acg Thr	ctt Leu 555	cct Pro	cat His	gca Ala	atg Met	att Ile 560	1680
tca Ser	ggt Gly	ggt Gly	gtt Val	tct Ser 565	aac Asn	gtc Val	tct Ser	ttt Phe	tct Ser 570	ttc Phe	cgt Arg	gga Gly	aat Asn	aac Asn 575	tac Tyr	1728
gtt Val	cgt Arg	gaa Glu	gct Ala 580	atc Ile	cat His	gcc	gta Val	ttt Phe 585	tta Leu	tat Tyr	cac His	tgt Cys	ttt Phe 590	.aaa Lys	aat Asn	1776
ggt Gly	atg Met	gat Asp 595	atg Met	ggc Gly	atc Ile	gta Val	aat Asn 600	gcg Ala	G17 999	cag Gln	ctg Leu	gaa Glu 605	ata Ile	tat Tyr	gat Asp	1824
aac Asn	gta Val 610	cca Pro	gaa Glu	gat Asp	ctg Leu	cgt Arg 615	gaa Glu	gcg Ala	gtt Val	gaa Glu	gat Asp 620	gtg Val	gta Val	ttg Leu	aat Asn	1872
cgt Arg 625	cga Arg	gat Asp	gat Asp	tct Ser	acg Thr 630	gag Glu	cgt Arg	tta Leu	ctt Leu	gat Asp 635	att Ile	gca Ala	act Thr	gag Glu	tat Tyr 640	1920
tta Leu	gaa Glu	cga Arg	gct Ala	gtt Val 645	ggt Gly	aaa Lys	gtt Val	gaa Glu	gat Asp 650	aaa Lys	tct Ser	gct Ala	tta Leu	gag Glu 655	tgg Trp	1968
cgt Arg	gac Asp	tgg Trp	cct Pro 660	gtt Val	gaa Glu	aaa Lys	cgt Arg	ctt Leu 665	gag Glu	cat His	tct Ser	cta Leu	gtg Val 670	aag Lys	G1 y 999	2016
ata Ile	aca Thr	gag Glu 675	ttt Phe	att Ile	gtc Val	gaa Glu	gat Asp 680	aca Thr	gaa Glu	gaa Glu	gca Ala	cga Arg 685	atc Ile	aat Asn	gca Ala	2064
Glu	aga Arg 690	cca Pro	ata Ile	gag Glu	gta Val	att Ile 695	gaa Glu	ggg ggg	cca Pro	ttg Leu	atg Met 700	gac Asp	gga Gly	atg Met	aac Asn	2112
gtc Val 705	gtt Val	ggt Gly	gat Asp	Leu	ttt Phe 710	999 Gly	gaa Glu	gga Gly	aaa Lys	atg Met 715	ttc Phe	ctt Leu	ccc Pro	caa Gln	gta Val 720	2160
gta Val	aag Lys	tct Ser	Ala	cgt Arg 725	gta Val	atg Met	aaa Lys	caa Gln	gct Ala 730	gtt Val	gct Ala	cat His	tta Leu	gaa Glu 735	ccg Pro	2208

tti Phe	atte Ile	aat Asr	gcg Ala 740	Ser	: aaa : Lys	ı gaa Glu	gtt Val	99a Gly 745	' Ala	aca Thr	aac Asn	ggt Gly	aaa Lys 750	Ile	ctt Leu	2256
tta Lev	gca Ala	aca Thr 755	Val	aaa Lys	ggt Gly	gat Asp	gtt Val 760	. His	gat Asp	att	ggt Gly	aag Lys 765	Asn	atc Ile	gtt Val	2304
ggg	gtg Val 770	Val	tta Leu	cag Gln	tgt Cys	aat Asn 775	aac Asn	tat Tyr	gaa Glu	ata Ile	att Ile 780	Asp	ctt Leu	ggt	gtc Val	2352
ме с 785	Val	ser	Сув	Glu	790		Leu	Lys	Val	Ala 795	Lys	Glu	Glu	Asn	Val 800	2400
Asp	Ile	Ile	Gly	Leu 805	Ser	gga Gly	Leu	Ile	Thr 810	Pro	Ser	Leu	Asp	Glu 815	Met	2448
Val	His	Val	820	Lys	Glu	atg Met	Glu	Arg 825	Gln	Gly	Phe	Ąap	Leu 830	Pro	Leu	2496
Leu	Ile	Gly 835	Gly	Ala	Thr	act Thr	Ser 840	Lys	Ala	His	Thr	Ala 845	Val	Lys	Ile	2544
Glu	G1n 850	Asn	Tyr	Ser	Gln	cct Pro 855	Val	Val	Tyr	Val	Asn 860	Asn	Ala	Ser	Arg	2592
865	Val	Gly	Val	Сув	Thr 870	tca Ser	Leu	Leu	Ser	Asn 875	Glu	Leu	Lys	Pro	Ser 880	2640
Phe	Val	Glu	Lув	Leu 885	Asp	att Ile	Asp	Tyr	Glu 890	Arg	Val	Arg	Glu	Gln 895	His	2688
Ser	Arg	Lys	900	Pro	Arg	act Thr	Lys	Pro 905	Val	Thr	Leu	Glu	Val 910	Ala	Arg	2736
Ala	Asn	Lys 915	Val	Ala	Ile	gac Asp	Trp 920	Ala	Ser	Tyr	Thr	Pro 925	Pro	Val	Pro	2784
Leu	1930 930	Pro	Gly	Val	His	ata Ile 935	Phe	Asp	Asn	Phe	Asp 940	Val	Ser	Thr	Leu	2832
Arg 945	Asn	Tyr	Ile .	Asp	Trp 950	acc Thr	Pro	Phe	Phe	Met 955	Thr	Trp	Ser	Leu	Val 960	2880
gga Gly	aaa Lys	tac Tyr	Pro :	aag Lys 965	atc Ile	tta Leu	gag Glu	His	gaa Glu 970	gaa Glu	gtt Val	ggt Gly	gaa Glu	gaa Glu 975	gcc Ala	2928
aaa	cga	tta	ttt	aaa	gat	gca .	aat	gat	cta	tta	gat	cga	gtt	gaa	aaa	2976

7		_		_		_				203						
			300	•				989	5				990)	Lys	
	Oly	995	n neu	. Dye	, wid	Arg	1000	met)	c Cyi	B Ala	Leu	Phe 1005	Pro	Ala	tcc Ser	3024
1	010	GIY	Авр	Asp) TIE	GIU 1015	Val	Туз	Thi	r Asp	Glu 1020	Ser	Arg	Thr	aca Thr	3072
1025		Dyb	val	neu	103	ASI 0	ren	Arg	, Glr	1 Gln 103		Glu	Lys	Pro	Lys 1040	3120
Gly	PHE	ABII	Tyr	1045	тел	ser	Asp	Тух	1050	: Ala	ccc Pro	Lys	Glu	Ser 1055	Gly	3168
2,0		, J	1060	116	GIY	GIA	Pne	A1a 1065	Val	Thr	ggt Gly	Gly	11e 1070	Gly	Glu	3216
arg .	1	075	MIG	wab	GIU	Tyr 1	080 TAB	Ala	Asn	Gly		Asp .085	Tyr	Asn	Ala	3264
110	090	116	GIN	ATA	vai	095	Asp	Arg	Leu	Ala	gaa Glu l100	Ala	Phe	Ala	Glu	3312
tat (Tyr) 1105	beu i	ulb	GIU	гув	va1 1110	Arg	Lys	Glu	Ile	Trp 1115	Gly	Tyr	Ser	Pro	Asn 1120	3360
gag a Glu 1	1114. 1	beu	3er 1	125	Asp .	Asp	Leu	Ile 1	Arg 130	Glu	Lys	Tyr	Gln 1	Gly 135	Ile	3408
cgt o	10 4	1	140	GIŸ	ıyr .	Pro .	Ala 1	Сув 145	Pro	Glu	His '	Thr 1	Glu 150	Lys	Gly	3456
gct t Ala L	11	.55	31U .	Leu	met 1	1:	160	GIU	Glu	Ser	Ile (31y : 165	Met	Ser	Leu	3504
aca t Thr S 11	70	er	ıyr 1	Ala I	Met 1	175 175	Pro (Gly	Ala	Ser 1	Val 9	Ser (Gly I	Met	Tyr	3552
ttt t Phe S 1185	er n	19 1	10 1	чер :	ser <i>F</i> 1190	arg 1	yr 1	Phe .	Ala	Ile 1 1195	Ala G	Sln :	Ile (Gln	Gln 1200	3600
gat c	aa g ln A	cc g la G	niu s	er 1 105	tat g Tyr A	icc g la A	at d	urg !	aaa Lys 210	ggt (Gly :	tgg a Irp A	at a sn 1	let 1	ctt (Leu (215	gaa 31u	3648
gct ga Ala G	ag aa lu L	ys T	gg t rp L 20	ta g	gt c Sly P	ca a	sn I	tg a eu i 25	aat Asn	taa						3681

<210> 46 <211> 1226 <212> PRT

<213> Vibrio fisheri

<400> 46

Val Ala Gly Ser Asn Ile Lys Val Gln Ile Glu Lys Gln Leu Ser Glu

1 5 10 15

Arg Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Gly Tyr
20 25 30

Lys Phe Glu Glu Lys Asp Tyr Arg Gly Gly Arg Phe Asn Gln Trp His 35 40

Cys Asp Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Ser Gln Pro Gln 50 55 60

Ile Ile Arg Asp Ile His Glu Ala Tyr Leu Glu Ala Gly Ala Asp Ile
65 70 75 80

Leu Glu Thr Asn Thr Phe Asn Ala Thr Thr Ile Ala Met Ala Asp Tyr

Asp Met Glu Ser Leu Ser Glu Glu Ile Asn Phe Glu Ala Ala Lys Leu 100 105 110

Ala Arg Glu Val Ala Asp Lys Trp Thr Glu Lys Thr Pro Asn Lys Pro 115 120 125

Arg Tyr Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Cys Ser Ile 130 135 140

Ser Pro Asp Val Asn Asp Pro Gly Phe Arg Asn Val Ser Phe Asp Glu 145 150 155 160

Leu Val Glu Ala Tyr Ser Glu Ser Thr Arg Ala Leu Ile Arg Gly Gly 165 170 175

Ser Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys 180 185 190

Ala Cys Ser Phe Ala Val Glu Ser Val Phe Glu Glu Leu Gly Ile Thr
195 200 205

Leu Pro Val Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr 210 225 220

Leu Ser Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ala Leu Arg His Val 225 230 235 240

Lys Pro Ile Ser Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu 245 250 255

Arg Glu Tyr Val Ser Glu Leu Ser Arg Ile Ser Glu Cys Tyr Val Ser 260 265 270

Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu 275 280 285

205

Ser Pro Glu Asp Met Ala Glu His Val Ala Glu Trp Ala Ser Ser Gly 290 295 300

Phe Leu Asn Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile 305 310 315 320

Arg Gln Met Ala Leu Val Val Glu Gly Val Lys Pro Arg Gln Leu Pro 325 330 335

Glu Leu Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Thr Ile 340 345 350

Glu Lys Asp Ser Leu Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr 355 360 365

Gly Ser Ala Arg Phe Lys Arg Leu Ile Lys Glu Glu Leu Tyr Asp Glu 370 375 380

Ala Leu Ser Val Ala Gln Glu Gln Val Glu Asn Gly Ala Gln Ile Ile 385 390 395 400

Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Cys Met Val 405 410 415

Arg Phe Leu Asn Leu Cys Ala Ser Glu Pro Glu Ile Ser Lys Val Pro
420 425 430

Val Met Val Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys 435 440 445

Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Leu Lys Glu Gly
450 460

Lys Glu Lys Phe Val His Gln Ala Lys Leu Ile Arg Arg Tyr Gly Ala 465 470 480

Ala Val Ile Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Arg 485 490 495

Glu Arg Lys Ile Glu Ile Cys Thr Asn Ala Tyr Asn Ile Leu Val Asp
500 505 510

Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe 515 520 525

Ala Val Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Val Asp Phe 530 535 540

Ile Glu Ala Val Gly Asp Ile Lys Arg Thr Leu Pro His Ala Met Ile 545 550 555 560

Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn Tyr
565 570 575

Val Arg Glu Ala Ile His Ala Val Phe Leu Tyr His Cys Phe Lys Asn 580 585 590

Gly Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr Asp 595 600 605

Asn Val Pro Glu Asp Leu Arg Glu Ala Val Glu Asp Val Val Leu Asn 610 620

- Arg Arg Asp Asp Ser Thr Glu Arg Leu Leu Asp Ile Ala Thr Glu Tyr 625 630 635 640
- Leu Glu Arg Ala Val Gly Lys Val Glu Asp Lys Ser Ala Leu Glu Trp
 645 650 655
- Arg Asp Trp Pro Val Glu Lys Arg Leu Glu His Ser Leu Val Lys Gly
 660 665 670
- Ile Thr Glu Phe Ile Val Glu Asp Thr Glu Glu Ala Arg Ile Asn Ala 675 680 685
- Glu Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn 690 700
- Val Val Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val 705 710 715 720
- Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Glu Pro 725 730 735
- Phe Ile Asn Ala Ser Lys Glu Val Gly Ala Thr Asn Gly Lys Ile Leu 740 745 750
- Leu Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val 755 760 765
- Gly Val Val Leu Gln Cys Asn Asn Tyr Glu Ile Ile Asp Leu Gly Val 770 780
- Met Val Ser Cys Glu Thr Ile Leu Lys Val Ala Lys Glu Glu Asn Val 785 790 795 800
- Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met 805 810 815
- Val His Val Ala Lys Glu Met Glu Arg Gln Gly Phe Asp Leu Pro Leu 820 825 830
- Leu Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile 835 840 845
- Glu Gln Asn Tyr Ser Gln Pro Val Val Tyr Val Asn Asn Ala Ser Arg 850 855 860
- Ala Val Gly Val Cys Thr Ser Leu Leu Ser Asn Glu Leu Lys Pro Ser 865 870 875 880
- Phe Val Glu Lys Leu Asp Ile Asp Tyr Glu Arg Val Arg Glu Gln His 885 890 895
- Ser Arg Lys Gln Pro Arg Thr Lys Pro Val Thr Leu Glu Val Ala Arg 900 905 910
- Ala Asn Lys Val Ala Ile Asp Trp Ala Ser Tyr Thr Pro Pro Val Pro 915 920 925
- Leu Lys Pro Gly Val His Ile Phe Asp Asn Phe Asp Val Ser Thr Leu 930 935 940
- Arg Asn Tyr Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Val

945 950 955

960

Gly Lys Tyr Pro Lys Ile Leu Glu His Glu Glu Val Gly Glu Glu Ala 965 970 975

Lys Arg Leu Phe Lys Asp Ala Asn Asp Leu Leu Asp Arg Val Glu Lys 980 985 990

Glu Gly Leu Leu Lys Ala Arg Gly Met Cys Ala Leu Phe Pro Ala Ser 995 1000 1005

Ser Val Gly Asp Asp Ile Glu Val Tyr Thr Asp Glu Ser Arg Thr Thr 1010 1015 1020

Val Ala Lys Val Leu His Asn Leu Arg Gln Gln Thr Glu Lys Pro Lys 1025 1030 1035 1040

Gly Phe Asn Tyr Cys Leu Ser Asp Tyr Ile Ala Pro Lys Glu Ser Gly 1045 1050 1055

Lys Asn Asp Trp Ile Gly Gly Phe Ala Val Thr Gly Gly Ile Gly Glu 1060 1065 1070

Arg Glu Leu Ala Asp Glu Tyr Lys Ala Asn Gly Asp Asp Tyr Asn Ala 1075 1080 1085

Ile Met Ile Gln Ala Val Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1090 1095 1100

Tyr Leu His Glu Lys Val Arg Lys Glu Ile Trp Gly Tyr Ser Pro Asn 1105 1110 1115 1120

Glu Thr Leu Ser Asn Asp Asp Leu Ile Arg Glu Lys Tyr Gln Gly Ile 1125 1130 1135

Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Gly
1140 1145 1150

Ala Leu Trp Glu Leu Met Asn Val Glu Glu Ser Ile Gly Met Ser Leu 1155 1160 1165

Thr Ser Ser Tyr Ala Met Trp Pro Gly Ala Ser Val Ser Gly Met Tyr 1170 1175 1180

Phe Ser His Pro Asp Ser Arg Tyr Phe Ala Ile Ala Gln Ile Gln Gln 1185 1190 1195 1200

Asp Gln Ala Glu Ser Tyr Ala Asp Arg Lys Gly Trp Asn Met Leu Glu 1205 1210 1215

Ala Glu Lys Trp Leu Gly Pro Asn Leu Asn 1220 1225

<210> 47

<211> 3780

<212> DNA

<213> Agrobacterium tumefaciens

<220>

<221> CDS

<222> (1) .. (3777)

<223> 15887359

<40	0> 4	7														
gtg Val 1	Pro	gtg Val	ttt Phe	gac Asp) Asp	ctg Leu	ttt Phe	ggc Gly	cct Pro) Glu	999 Gly	gca Ala	aag Lys	Arg	_	48
Gly	gcg	gaa Glu	att Ile 20	t t c Phe	aag Lys	gcg Ala	ttg Lev	cgc Arg 25	Asp	gcc Ala	gcc	agc Ser	gaa Glu 3,0	Arg	atc Ile	96
ctc Leu	att Ile	ctc Leu 35	Asp	ggt	gcc Ala	atg Met	ggc Gly 40	Thr	Glr.	atc Ile	cag Gln	ggt Gly 45	ctc Leu	ggt Gly	ttt Phe	144
Asp	Glu 50	Asp	His	Phe	cgt Arg	Gly 55	Asp	Arg	Phe	lle	Gly 60	Сув	Ala	Сув	His	192
Gln 65	Lys	Gly	Asn	Asn	gac Asp 70	Leu	Leu	Ile	Leu	Thr 75	Gln	Pro	Asp	Āla	Ile 80	240
Glu	Glu	Ile	His	Tyr 85	cgc Arg	Tyr	Ala	Met	Ala 90	Gly	Ala	Asp	Ile	Leu . 95	Glu	288
Thr	Asn	Thr	Phe 100	Ser	tcc Ser	Thr	Arg	Ile 105	Ala	Gln	Ala	Asp	Tyr 110	Glu	Met	336
Glu	Asn	115	Val	Tyr	gat Asp	Leu	Asn 120	Arg	Glu	Gly	Ala	Ala 125	Ile	Val	Arg	384
Arg	130	Ala	Gln	Arg	gcc Ala	Glu 135	Arg	Glu	Asp	Gly	Arg 140	Arg	Arg	Phe	Val	432
145	Gly	Ala	Ile	Gly	ccg Pro 150	Thr	Asn	Arg	Thr	Ala 155	Ser	Ile	Ser	Pro	Asp 160	480
Val	Asn	Asn	Pro	Gly 165	tac Tyr	Arg	Ala	Val	Ser 170	Phe	Asp	Asp	Leu	Arg 175	Ile	528
Ala	Tyr	Gly	Glu 180	Gln	atc Ile	Asp	Gly	Leu 185	Ile	Ąsp	Gly	Gly	Ala 190	Asp	Ile	576
Ile	Leu	11e 195	Glu	Thr	atc Ile	Phe	Asp 200	Thr	Leu	Asn	Ala	L ув 205	Ala	Ala	Ile	624
Pne .	210	Cys	Glu	Glu		Phe 215	Glu	Ala	Lys	Gly	11e 220	Arg	Leu	Pro	Val	672
atg Met 225	atc Ile	tca Ser	ggc Gly	Thr	atc Ile 230	acc Thr	gac Asp	ctt Leu	tcc Ser	ggt Gly 235	cgc Arg	acg Thr	ttg Leu	tcc Ser	ggc Gly 240	720

cag Gln	acg Thr	cct Pro	tcg Ser	gcg Ala 245	ttc Phe	tgg Trp	aac Asn	tcg Ser	gtg Val 250	cgc	cac His	gcc Ala	aac Asn	ccc Pro 255	ttc Phe	768
acc Thr	atc Ile	ggc	ctc Leu 260	aac Asn	tgc Cys	gcg Ala	ctc Leu	ggt Gly 265	gcg Ala	gat Asp	gcc	atg Met	cgc Arg 270	ccg Pro	cat His	816
ctg Leu	cag Gln	gaa Glu 275	Leu	tcc Ser	gat Asp	gtg Val	gcc Ala 280	gac Asp	acc Thr	ttt Phe	gtc Val	tgc Cys 285	gcc Ala	tat Tyr	ccg Pro	864
aat Asn	gcc Ala 290	ggc	ctg Leu	ccg	aac Asn	gag Glu 295	ttc Phe	ggc Gly	caa Gln	tat Tyr	gac Asp 300	gaa Glu	acg Thr	ccc Pro	gag Glu	912
atg Met 305	atg Met	gcg Ala	cgc Arg	cag Gln	gtt Val 310	gag Glu	ggc	ttc Phe	gtt Val	cgt Arg 315	gac Asp	ggt Gly	ctc Leu	gtc Val	aac Asn 320	960
atc Ile	gtc Val	ggc	ggt Gly	tgc Cys 325	tgc Cys	ggt Gly	tcg Ser	acg Thr	ccg Pro 330	Glu	cat His	atc Ile	cgg Arg	gcg Ala 335	att Ile	1008
gcc Ala	gaa Glu	gcc Ala	gtc Val 340	aag Lys	gat Asp	tac Tyr	aag Lys	ccc Pro 345	cgc Arg	gaa Glu	att Ile	cct Pro	gaa Glu 350	cac Ris	aag Lys	1056
ccg Pro	ttc Phe	atg Met 355	tcg Ser	ctt Leu	tcc Ser	ggc Gly	ctt Leu 360	gaa Glu	ccc Pro	ttc Phe	gtg Val	ctg Leu 365	acc Thr	aag Lys	gac Asp	1104
att Ile	ccc Pro 370	ttc Phe	gtc Val	aac Asn	gtg Val	ggc Gly 375	ga g Glu	cgc Arg	acc Thr	aac Asn	gtc Val 380	acc Thr	ggt Gly	tcg Ser	gcc Ala	1152
Arg 385	Phe	Arg	Lys	Leu	atc Ile 390	Thr	Ala	Gly	Asp	Tyr 395	Thr	Ala	Ala	Leu	Ala 400	1200
gtt Val	gcc Ala	cgc Arg	gac Asp	cag Gln 405	gtg Val	gaa Glu	aac Asn	ggc Gly	gcg Ala 410	cag Gln	atc Ile	atc Ile	Asp gac	atc Ile 415	aac Asn	1248
atg Met	gat Asp	gag Glu	ggc Gly 420	ctg Leu	atc Ile	gat Asp	tcg Ser	gaa Glu 425	aag Lys	gcg Ala	atg Met	gtc Val	gag Glu 430	ttc Phe	ctg Leu	1296
aac Asn	Leu	atc Ile 435	gcc Ala	gcc Ala	gag Glu	Pro	gac Asp 440	att Ile	gcc Ala	cgt Arg	gtg Val	ccc Pro 445	gtc Val	atg Met	atc Ile	1344
gac Asp	tca Ser 450	tcc Ser	aag Lys	ttc Phe	gag Glu	atc Ile 455	atc Ile	gag Glu	gcc Ala	ggc Gly	ctg Leu 460	aaa Lys	tgc Cys	gtg Val	cag Gln	1392
ggc Gly 465	aaa Lys	tcg Ser	atc Ile	Val	aat Asn 470	tcc Ser	att Ile	tcg Ser	ctg Leu	aag Lys 475	gaa Glu	ggc Gly	gag Glu	gag Glu	aag Lys 480	1440
ttt	ctc	cag	cag	gct	cgg	ctc	gtc	cac	aat	tac	ggt	gcg	gcg	gtt	gtc	1488

Phe	Leu	Gln	Gln	Ala 485	Arg	Leu	Val	His	Asn 490		Gly	Ala		Val 495	Val	
gtc Val	atg Met	gcc Ala	ttt Phe 500	Ąsp	gag Glu	gtc Val	61A 888	cag Gln 505	gcg Ala	gat Asp	acc Thr	tat Tyr	cag Gln 510	. cgc Arg	aag Lys	1536
gtg Val	gaa Glu	atc Ile 515	Сув	gcg Ala	cgc Arg	gcc Ala	tac Tyr 520	aag Lys	ctt Leu	ctg Leu	acc Thr	gaa Glu 525	aag Lys	gcc Ala	ggt Gly	1584
ctg Leu	tct Ser 530	ccg Pro	gaa Glu	gac Asp	atc Ile	atc Ile 535	ttc Phe	gac Asp	ccg Pro	aat Asn	gtg Val 540	ttt Phe	gcg Ala	gta Val	gct Ala	1632
acg Thr 545	ggc Gly	atc Ile	gag Glu	gag Glu	cac His 550	aat Asn	aat Asn	tac Tyr	ggc Gly	gtg Val 555	gac	ttc Phe	atc Ile	gag Glu	gcc Ala 560	1680
acc Thr	aag Lys	acc Thr	atc Ile	cgc Arg 565	gaa Glu	acc Thr	atg Met	ccg Pro	ctc Leu 570	acg Thr	cat His	att Ile	tcc Ser	999 Gly 575	ggc Gly	1728
gtt Val	tcc Ser	aac Asn	ctg Leu 580	tcc Ser	ttc Phe	tcc Ser	ttc Phe	cgc Arg 585	ggc	aat Asn	gag Glu	Pro	gtg Val 590	Arg	gag Glu	1776
Ala	Met	His 595	Ala	Val	Phe	Leu	Tyr 600	cac His	Ala	Ile	Gln	Val 605	Gly	Met	Asp	1824
Met	Gly 610	Ile	Val	Asn	Ala	Gly 615	Gln	ctt Leu	Ala	Val	Tyr 620	Авр	Asn	Ile	Asp	1872
Ala 625	Glu	Leu	Arg	Glu	Ala 630	Сув	Glu	gac Asp	Val	Val 635	Leu	Asn	Arg	Arg	Asp 640	1920
gat Asp	gcc Ala	acg Thr	gag Glu	cgt Arg 645	ctg Leu	ctc Leu	gag Glu	gtg Val	gcg Ala 650	gag Glu	cgt Arg	ttc Phe	cgt Arg	ggt Gly 655	acg Thr	1968
Gly	Glu	Lys	Gln 660	Ala	Lys	Val	Gln	gat Asp 665	Leu	Ser	Trp	Arg	Glu 670	Tyr	Pro	2016
gtt Val	gaa Glu	aag Lys 675	cgg Arg	ctg Leu	gaa Glu	cat His	gct Ala 680	ctg Leu	gtc Val	aac Asn	ggc Gly	att Ile 685	acc Thr	gac Asp	tat Tyr	2064
atc Ile	gag Glu 690	gcc Ala	gat Asp	acg Thr	Glu	gag Glu 695	gca Ala	cgc Arg	cag Gln	cag Gln	gcc Ala 700	gcc Ala	cgc Arg	ccg Pro	ctg Leu	2112
cat His 705	gtc Val	atc Ile	gaa Glu	G1 y 999	ccg Pro 710	ctg Leu	atg Met	gcc Ala	ggt Gly	atg Met 715	aat Asn	gtg Val	gtg Val	ggt Gly	gac Asp 720	2160
ctg Leu	ttc Phe	ggt Gly	Ser	ggc Gly 725	aag Lys	atg Met	ttc Phe	ctg Leu	cca Pro 730	cag Gln	gtg Val	gtg Val	aaa Lys	tcc Ser 735	gcc Ala	2208

Arg	, val	. Met	740	Glr	Ala	val	. Ala	745	Leu	. Leu	Pro	Tyr	750	Glu	gag Glu	2256
gaa Glu	aag Lys	r cgc Arg 755	Leu	, aat Asn	ggc	ggt Gly	Ser 760	Glu	Arg	agt Ser	gcc Ala	gcc Ala 765	Gly	aag Lys	gtg Val	2304
Leu	770	Ala	Thr	· Val	Lys	775	Asp	Val	His	qaA	780	Gly	ГÀв	Asn	atc Ile	2352
Val 785	Gly	Val	Val	cta Leu	Ala 790	Сув	Asn	Asn	Tyr	Glu 795	Ile	Ile	Asp	Leu	Gly 800	2400
val	Met	Val	Pro	acg Thr 805	Thr	Lys	Ile	Leu	Glu 810	Thr	Ala	Ile	Ala	Glu 815	Lys	2448
Val	Asp	Val	11e 820	ggc	Leu	Ser	Gly	Leu 825	Ile	Thr	Pro	Ser	Leu 830	Āsp	Glu	2496
Met	Val	835	Val	gcg Ala	Ala	Glu	Met 840	Glu	Arg	Gln	Gly	Phe 845	Ąap	Ile	Pro	2544
Leu	850	Ile	Gly	ggt Gly	Ala	Thr 855	Thr	Ser	Arg	Val	His 860	Thr	Ala	Val	Lys	2592
11e 865	His	Pro	Arg	tac Tyr	Glu 870	Gln	Gly	Gln	Ala	11e 875	Tyr	Val	Thr	Asp	Ala 880	2640
Ser	Arg	Ala	Val	ggc Gly 885	Val	Val	Ser	Ala	Leu 890	Leu	Ser	Glu	Glu	Gln 895	Lys	2688
Pro	Ala	Tyr	900	gac Asp	Gly	Ile	Arg	Ala 905	Glu	Tyr	Ala	Lys	Val 910	Ala	Glu	2736
Ala	His	Ala 915	Arg	aat Asn	Glu	Arg	Glu 920	Lys	Gln	Arg	Leu	Pro 925	Leu	Ser	Arg	2784
	Arg 930	Glu	Asn	Ala	His	Lys 935	Ile	Asp	Trp	Ser	Ser 940	Tyr	Ser	Val	Val	2832
Lys 945	Pro	GIn	Phe		Gly 950	Thr	Lys	Val	Phe	Glu 955	Thr	Tyr	Asp	Leu	Glu 960	2880
gag Glu	Leu	ser .	Arg	Tyr 965	Ile .	qaA	Trp	Thr	Pro 970	Phe	Phe	Gln	Thr	Trp 975	Glu	2928
ttg	aag	ggc	cgt	ttc	ccg	gcg	atc	ctt	gaa	gac	gaa	aag	cag	ggc	gag	2976

									2	12						
Leu	Lys	Gly	Arg 980	Phe	Pro	Ala	Ile	Leu 985	Glu	Asp	Glu	Lys	Gln 990	Gly	Glu	
gcg Ala	gcg Ala	cgg Arg 995	cag Gln	ctt Leu	tat Tyr	Ala	gat Asp 1000	Ala	cag Gln	gcc Ala	Met	ctt Leu 1005	gcg Ala	aag Lys	atc Ile	3024
Ile	gag Glu 1010	gaa Glu	aag Lys	tgg Trp	Phe	cga Arg 1015	cca Pro	cgc Arg	gcg Ala	Val	atc Ile 1020	ggc Gly	ttc Phe	tgg Trp	ccg Pro	3072
gcc Ala 1025	Asn	gcc Ala	gtg Val	ggt Gly	gac Asp 103	Asp	atc Ile	agg Arg	ctc Leu	ttt Phe 1035	Thr	gat Asp	gaa Glu	ggt Gly	cgg Arg 1040	3120
aag Lys	gaa Glu	gag Glu	ttg Leu	gcg Ala 045	acg Thr	ttc Phe	ttc Phe	Thr	ctg Leu 050	cgc Arg	cag Gln	cag Gln	Leu	tcc Ser 1055	Lys	3168
Arg	Asp	Gly	cgt Arg 1060	Pro	neA	Val	Ala	Leu 1065	Ser	Asp	Phe	Val	Ala 1070	Pro	Val	3216
Asp	Ser 1	Gly 1075	_gtt Val	Ala	Asp	Tyr	Val 1080	Gly	Gly	Phe	Val	Val 1085	Thr	Ala	Gly	3264
Ile	gag Glu .090	gaa Glu	gtg Val	gcg Ala	Ile	gcc Ala 1095	gag Glu	cgc Arg	ttc Phe	Glu	cgg Arg .100	gcc Ala	aat Asn	gac Asp	gat Asp	3312
tat Tyr 1105	Ser	tcc Ser	atc Ile	ctc Leu	gtc Val 1110	Lys	gcg Ala	ttg Leu	gct Ala	gac Asp 1115	Arg	ttt Phe	gcc Ala	gaa Glu	gcc Ala 1120	3360
ttt Phe	gcc Ala	gag Glu	Arg	atg Met 125	cat His	gag Glu	cgc Arg	Val	cgc Arg 130	aag Lys	gag Glu	ttc Phe	Trp	ggt Gly 1135	tat Tyr	3408
gcg Ala	ccg Pro	Asp	gag Glu 140	gct Ala	ctt Leu	gcc Ala	Gly	gac Asp 145	gat Asp	ctg Leu	ata Ile	Gly	gaa Glu 150	gcc Ala	tat Tyr	3456
gcc Ala	Gly	atc Ile 155	cgc Arg	ccg Pro	gca Ala	Pro	ggt Gly .160	tat Tyr	ccg Pro	gcc Ala	Gln	ccg Pro 165	gac Asp	cac His	acc Thr	3504
gaa Glu 1	aag Lys 170	aag Lys	acg Thr	ctg Leu	Phe	gct Ala 175	ctg Leu	ctg Leu	gac Asp	Ala	acc Thr 180	aat Asn	gcg Ala	gcg Ala	ggt Gly	3552
gtg Val 1185	Glu	ttg Leu	acg (Glu	agc Ser 1190	Tyr	gcg Ala	atg Met	Trp	ccc Pro 1195	Gly	tcg Ser	tcg Ser	gtt Val	tcg Ser 1200	3600
ggc Gly	ctc Leu	tat Tyr	Ile	ggc Gly 205	cat His	ccc Pro	gaa Glu	Ser	tat Tyr 210	tat Tyr	ttc Phe	ggc Gly	Val	gcc Ala 215	aag Lys	3648
gtg Val	gag Glu	Arg .	gat (Asp (220	cag 31n	gtt Val	ctc Leu	qaA	tat Tyr 225	gcg Ala	cgc Arg	cgc Arg	Lys	gat Asp 230	atg Met	ccg Pro	3696

gtc aca gag gtg gag cgc tgg ctc ggg ccg gtg ctc aac tac gtg ccg 3744 Val Thr Glu Val Glu Arg Trp Leu Gly Pro Val Leu Asn Tyr Val Pro 1235 1240 1245

acc aac ggc gag gag aaa atc gac agc gct gcg tga Thr Asn Gly Glu Glu Lys Ile Asp Ser Ala Ala 1250 1255

3780

<210> 48 <211> 1259 <212> PRT

<213> Agrobacterium tumefaciens

<400> 48

Val Pro Val Phe Asp Asp Leu Phe Gly Pro Glu Gly Ala Lys Arg Asp
1 5 10 15

Gly Ala Glu Ile Phe Lys Ala Leu Arg Asp Ala Ala Ser Glu Arg Ile 20 25 30

Leu Ile Leu Asp Gly Ala Met Gly Thr Gln Ile Gln Gly Leu Gly Phe 35 40

Asp Glu Asp His Phe Arg Gly Asp Arg Phe Ile Gly Cys Ala Cys His 50 55

Gln Lys Gly Asn Asn Asp Leu Leu Ile Leu Thr Gln Pro Asp Ala Ile 65 70 75 80

Glu Glu Ile His Tyr Arg Tyr Ala Met Ala Gly Ala Asp Ile Leu Glu 85 90 95

Thr Asn Thr Phe Ser Ser Thr Arg Ile Ala Gln Ala Asp Tyr Glu Met
100 105 110

Glu Asn Ala Val Tyr Asp Leu Asn Arg Glu Gly Ala Ala Ile Val Arg 115 120 125

Arg Ala Ala Gln Arg Ala Glu Arg Glu Asp Gly Arg Arg Phe Val 130 135 140

Ala Gly Ala Ile Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp 145 150 155 160

Val Asn Asn Pro Gly Tyr Arg Ala Val Ser Phe Asp Asp Leu Arg Ile 165 170 175

Ala Tyr Gly Glu Gln Ile Asp Gly Leu Ile Asp Gly Gly Ala Asp Ile 180 185 190

Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile 195 200 205

Phe Ala Cys Glu Glu Arg Phe Glu Ala Lys Gly Ile Arg Leu Pro Val 210 215 220

Met Ile Ser Gly Thr Ile Thr Asp Leu Ser Gly Arg Thr Leu Ser Gly 225 230 235 240

Gln Thr Pro Ser Ala Phe Trp Asn Ser Val Arg His Ala Asn Pro Phe

245 250 255 Thr Ile Gly Leu Asn Cys Ala Leu Gly Ala Asp Ala Met Arg Pro His 265 Leu Gln Glu Leu Ser Asp Val Ala Asp Thr Phe Val Cys Ala Tyr Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Gln Tyr Asp Glu Thr Pro Glu 295 Met Met Ala Arg Gln Val Glu Gly Phe Val Arg Asp Gly Leu Val Asn 310 Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Glu His Ile Arg Ala Ile 330 Ala Glu Ala Val Lys Asp Tyr Lys Pro Arg Glu Ile Pro Glu His Lys Pro Phe Met Ser Leu Ser Gly Leu Glu Pro Phe Val Leu Thr Lys Asp Ile Pro Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Arg Phe Arg Lys Leu Ile Thr Ala Gly Asp Tyr Thr Ala Ala Leu Ala Val Ala Arg Asp Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile Asn 410 Met Asp Glu Gly Leu Ile Asp Ser Glu Lys Ala Met Val Glu Phe Leu 420 425 Asn Leu Ile Ala Ala Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Phe Glu Ile Ile Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu Lys 465 470 475 Phe Leu Gln Gln Ala Arg Leu Val His Asn Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Tyr Gln Arg Lys Val Glu Ile Cys Ala Arg Ala Tyr Lys Leu Leu Thr Glu Lys Ala Gly Leu Ser Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Gly Val Asp Phe Ile Glu Ala Thr Lys Thr Ile Arg Glu Thr Met Pro Leu Thr His Ile Ser Gly Gly 565 570

Val Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu 580 585 590

Ala Met His Ala Val Phe Leu Tyr His Ala Ile Gln Val Gly Met Asp 595 600 605

Met Gly Ile Val Asn Ala Gly Gln Leu Ala Val Tyr Asp Asn Ile Asp 610 615 620

Ala Glu Leu Arg Glu Ala Cys Glu Asp Val Val Leu Asn Arg Arg Asp 625 630 635 640

Asp Ala Thr Glu Arg Leu Leu Glu Val Ala Glu Arg Phe Arg Gly Thr 645 650 655

Gly Glu Lys Gln Ala Lys Val Gln Asp Leu Ser Trp Arg Glu Tyr Pro 660 665 670

Val Glu Lys Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Asp Tyr 675 680 685

Ile Glu Ala Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Leu 690 695 700

His Val Ile Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp 705 710 715 720

Leu Phe Gly Ser Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala 725 730 735

Arg Val Met Lys Gln Ala Val Ala Val Leu Leu Pro Tyr Met Glu Glu 740 745 750

Glu Lys Arg Leu Asn Gly Gly Ser Glu Arg Ser Ala Ala Gly Lys Val 755 760 765

Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile 770 780

Val Gly Val Val Leu Ala Cys Asn Asn Tyr Glu Ile Ile Asp Leu Gly
785 790 795 800

Val Met Val Pro Thr Thr Lys Ile Leu Glu Thr Ala Ile Ala Glu Lys 805 810 815

Val Asp Val Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu 820 825 830

Met Val His Val Ala Ala Glu Met Glu Arg Gln Gly Phe Asp Ile Pro 835 840 845

Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val Lys 850 855 860

Ile His Pro Arg Tyr Glu Gln Gly Gln Ala Ile Tyr Val Thr Asp Ala 865 870 875

Ser Arg Ala Val Gly Val Val Ser Ala Leu Leu Ser Glu Glu Gln Lys 885 890 895

Pro Ala Tyr Ile Asp Gly Ile Arg Ala Glu Tyr Ala Lys Val Ala Glu 900 905 910

- Ala His Ala Arg Asn Glu Arg Glu Lys Gln Arg Leu Pro Leu Ser Arg 915 920 925
- Ala Arg Glu Asn Ala His Lys Ile Asp Trp Ser Ser Tyr Ser Val Val 930 935 940
- Lys Pro Gln Phe Phe Gly Thr Lys Val Phe Glu Thr Tyr Asp Leu Glu 945 950 955 960
- Glu Leu Ser Arg Tyr Ile Asp Trp Thr Pro Phe Phe Gln Thr Trp Glu 965 970 975
- Leu Lys Gly Arg Phe Pro Ala Ile Leu Glu Asp Glu Lys Gln Gly Glu 980 985 990
- Ala Ala Arg Gln Leu Tyr Ala Asp Ala Gln Ala Met Leu Ala Lys Ile 995 1000 1005
- Ile Glu Glu Lys Trp Phe Arg Pro Arg Ala Val Ile Gly Phe Trp Pro 1010 1015 1020
- Ala Asn Ala Val Gly Asp Asp Ile Arg Leu Phe Thr Asp Glu Gly Arg 1025 1030 1035 1040
- Lys Glu Glu Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Leu Ser Lys 1045 1050 1055
- Arg Asp Gly Arg Pro Asn Val Ala Leu Ser Asp Phe Val Ala Pro Val 1060 1065 1070
- Asp Ser Gly Val Ala Asp Tyr Val Gly Gly Phe Val Val Thr Ala Gly 1075 1080 1085
- Ile Glu Glu Val Ala Ile Ala Glu Arg Phe Glu Arg Ala Asn Asp Asp 1090 1095 1100
- Tyr Ser Ser Ile Leu Val Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala 1105 1110 1115 1120
- Phe Ala Glu Arg Met His Glu Arg Val Arg Lys Glu Phe Trp Gly Tyr 1125 1130 1135
- Ala Pro Asp Glu Ala Leu Ala Gly Asp Asp Leu Ile Gly Glu Ala Tyr 1140 1145 1150
- Ala Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr 1155 1160 1165
- Glu Lys Lys Thr Leu Phe Ala Leu Leu Asp Ala Thr Asn Ala Ala Gly 1170 1175 1180
- Val Glu Leu Thr Glu Ser Tyr Ala Met Trp Pro Gly Ser Ser Val Ser 1185 1190 1195 1200
- Gly Leu Tyr Ile Gly His Pro Glu Ser Tyr Tyr Phe Gly Val Ala Lys 1205 1210 1215
- Val Glu Arg Asp Gln Val Leu Asp Tyr Ala Arg Arg Lys Asp Met Pro 1220 1225 1230
- Val Thr Glu Val Glu Arg Trp Leu Gly Pro Val Leu Asn Tyr Val Pro

1245

1240

1235

Thr Asn Gly Glu Glu Lys Ile Asp Ser Ala Ala 1255 <210> 49 <211> 2718 <212> DNA <213> Ralstonia solanacearum <220> <221> CDS <222> (1) .. (2715) <223> RSOL_GMI1000 atg acc gac cac ctc atg cgc ctc tcc ggc ctc gaa ccg ttc aac atc Met Thr Asp His Leu Met Arg Leu Ser Gly Leu Glu Pro Phe Asn Ile 5 10 gge gag gae aeg etg tte gte aac gte gge gaa ege ace aac gte ace 96 Gly Glu Asp Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr 20 25 gga tcc aag gcg ttc gcg cgc atg atc ctc aac agc cag ttc gac gag 144 Gly Ser Lys Ala Phe Ala Arg Met Ile Leu Asn Ser Gln Phe Asp Glu 35 40 geg etc gec gtg gea ege cag cag gtc gag aac ggc geg cag gtc atc 192 Ala Leu Ala Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile 50 55 gac atc aac atg gac gag gcc atg ctc gac tcc aag gcg gcg atg gtg 240 Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Lys Ala Ala Met Val 65 70 cgc ttc ctg aac ctg atc gcc tcg gag ccg gac atc gcg cgc gtg ccg 288 Arg Phe Leu Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Arg Val Pro 90 atc atg atc gac tcg tcc aag tgg gag gtg atc gag gcc ggc ctg aag 336 Ile Met Ile Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys tgc gtg cag ggc aag gcc atc gtc aac tcg atc tcg ctc aag gaa ggc 384 Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Glu Gly 120 gag gaa cag ttc gcc cac cac gcc aag ctg atc aag cgc tac ggc gcc 432 Glu Glu Gln Phe Ala His His Ala Lys Leu Ile Lys Arg Tyr Gly Ala 130 135 gcc gcc gtg gtg atg gcc ttc gac gag cag ggc cag gcc gac acg ttc 480 Ala Ala Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Phe

gcg cgc aag acc gag atc tgc aag cgc agc tat gac ttc ctc gtg aac Ala Arg Lys Thr Glu Ile Cys Lys Arg Ser Tyr Asp Phe Leu Val Asn

cag gtc ggc ttt gcg ccg gaa gac atc atc ttc gat ccg aac atc ttc

170

Gln	Val	Gly	Phe 180	Ala	Pro	Glu	Asp	Ile 185	Ile	Phe	Asp	Pro	Asn 190	Ile	Phe	
gcg Ala	gtc Val	gcc Ala 195	acc Thr	ggc Gly	atc Ile	gag Glu	gag Glu 200	cac His	aac Asn	aac Asn	tac Tyr	gcc Ala 205	gtc Val	gac Asp	ttc Phe	624
atc Ile	gag Glu 210	gcc Ala	acg Thr	cgc Arg	tgg Trp	atc Ile 215	aag Lys	cag Gln	aaa Lys	ttg Leu	ccg Pro 220	cac His	gcc Ala	aag Lys	gtg Val	672
agc Ser 225	ggc Gly	ggc	gtg Val	tcg Ser	aac Asn 230	gtc Val	tcg Ser	ttc Phe	tcg Ser	ttc Phe 235	cgc Arg	ggc Gly	aac Asn	gac Asp	gtg Val 240	720
Val	Arg	Glu	Ala	Ile 245	His	acc Thr	Val	Phe	Leu 250	Тут	His	Ala	Ile	Gly 255	Ala	768
Gly	Met	Asp	Met 260	Gly	Ile	gtc Val	Asn	Ala 265	Gly	Gln	Leu	Gly	Val 270	Tyr	Glu	816
Asn	Leu	Ala 275	Pro	Glu	Leu	cgc Arg	Glu 280	Arg	Val	Glu	Asp	Val 285	Val	Leu	Asn	864
Arg	Arg 290	Pro	Авр	Ala	Thr	gac Asp 295	Arg	Leu	Leu	Glu	11e 300	Āla	Asp	Arg	Tyr	912
Lys 305	Gly	Gly	Gly	Ala	Lys 310	cgc Arg	Glu	Glu	Asn	Leu 315	Ala	Trp	Arg	Gln	Glu 320	960
Pro	Val	Glu	Lys	Arg 325	Leu	gcc Ala	His	Ala	Leu 330	Val	His	Gly	Ile	Thr 335	Ąsp	1008
Tyr	Val	Val	Glu 340	Asp	Thr	gag Glu	Glu	Val 345	Arg	Gln	Lys	Ile	Phe 350	Ala	Ala	1056
Gly	Gly	Arg 355	Pro	Ile	Gln	gtg Val	Ile 360	Glu	Gly	Pro	Leu	Met 365	Āsp	ĞÎy	Met	1104
Asn	11e 370	Val	Gly	Asp	Leu	ttc Phe 375	Gly	Ala	Gly	Lys	Met 380	Phe	Leu	Pro	Gln	1152
Val 385	Val	Lys	Ser	Ala	Arg 390	gtg Val	Met	Lys	Gln	Ala 395	Val	Ala	His	Leu	Ile 400	1200
Pro	Phe	Ile	Glu	Glu 405	Glu	aag Lys	Arg	Gln	Ile 410	Ala	Ala	Ala	Gly	Gly 415	Asp	1248
gtg Val	cgc Arg	tcg Ser	cgc Arg 420	ggc Gly	aag Lys	atc Ile	gtc Val	atc Ile 425	gcc Ala	acc Thr	gtg Val	aag Lys	ggc Gly 430	gac Asp	gtg Val	1296

cac His	gac Asp	atc Ile 435	Gly	aag Lys	aac Asn	atc	gtc Val 440	Thr	gtc Val	gtg Val	ctc Leu	cag Gln 445	tgc Cys	aac Asn	aac Asn	1344
ttc Phe	gaa Glu 450	Val	gtg Val	aac Asn	atg Met	ggc Gly 455	gtg Val	atg Met	gtc Val	ccg Pro	tgc Cys 460	aac Asn	gag Glu	atc Ile	ctg Leu	1392
gcc Ala 465	aag Lys	gcg Ala	aag Lys	gtc Val	gag Glu 470	ggc Gly	gcg Ala	gac Asp	atc	atc Ile 475	ggc Gly	ctg Leu	tcg Ser	ggc	ctg Leu 480	1440
atc Ile	aca Thr	ccg	tcg Ser	ctg Leu 485	gaa Glu	gag Glu	atg Met	gcc Ala	tac Tyr 490	gtg Val	gcc Ala	tcc Ser	gag Glu	atg Met 495	cag Gln	1488
cgc Arg	Asp	gag Glu	tac Tyr 500	ttc Phe	cgc Arg	gtg Val	aag Lys	aag Lys 505	atc Ile	ccg Pro	ctg Leu	ctg Leu	atc Ile 510	ggt Gly	ggc Gly	1536
gcg Ala	acc Thr	acg Thr 515	Ser	cgc Arg	gtg Val	cac His	acc Thr 520	gcc Ala	gtg Val	aag Lys	atc Ile	gcg Ala 525	ccc Pro	aat Asn	tac Tyr	1584
gaa Glu	ggc Gly 530	ccg Pro	gtc Val	gtg Val	tac Tyr	gtg Val 535	ccc Pro	gac Asp	gcc Ala	tcg Ser	cgc Arg 540	tcg Ser	gtg Val	agc Ser	gtg Val	1632
Ala 545	Ser	Ser	Leu	Leu	Ser 550	gac Asp	Glu	Ala	Ala	Ala 555	Arg	Tyr	Ile	Glu	Glu 560	1680
ctg Leu	cac His	gcc Ala	yac Yab	tac Tyr 565	gac Asp	cgc Arg	atc Ile	cgc Arg	acc Thr 570	cag Gln	cac His	gcc Ala	agc Ser	aag Lys 575	aaa Lys	1728
gcc Ala	atg Met	ccg Pro	atg Met 580	gtg Val	tcg Ser	ctg Leu	gcc Ala	gcc Ala 585	gcg Ala	cgc Arg	gcc Ala	aac Asn	аад L ув 590	acc Thr	cgġ Arg	1776
atc Ile	Asp	tgg Trp 595	tcg Ser	aac Asn	tac Tyr	acg Thr	ccg Pro 600	ccc Pro	aag Lys	ccc Pro	aag Lys	ttc Phe 605	gtc Val	ggc	cgc Arg	1824
cgc Arg	gtg Val 610	ttc Phe	cgc Arg	aac Asn	tac Tyr	gac Asp 615	ctg Leu	aac Asn	gag Glu	ctc Leu	gcg Ala 620	cag Gln	tac Tyr	atc Ile	gac Asp	1872
tgg Trp 625	ggc Gly	ccg Pro	ttc Phe	ttc Phe	cag Gln 630	acg Thr	tgg Trp	gac Asp	ctg Leu	gcc Ala 635	ggc Gly	aaa Lys	ttc Phe	ccc Pro	gac Asp 640	1920
atc Ile	ctc Leu	aac Asn	gac Asp	gcg Ala 645	atc Ile	gtc Val	ggc Gly	gaa Glu	tcg Ser 650	gcc Ala	cgc Arg	cgc Arg	gtg Val	ttc Phe 655	tcc Ser	1968
gac Asp	ggc Gly	aag Lys	agc Ser 660	atg Met	ctc Leu	gcg Ala	cgc Arg	ctg Leu 665	atc Ile	gcc Ala	gga Gly	cgc Arg	tgg Trp 670	ctg Leu	acg Thr	2016
gcc	aac	ggc	gtg	atc	gcg	ctg	ctg	ccg	gcc	aac	acc	gtc	aac	gac	gac	2064

220

									2	20						
Ala	Asn	Gly 675	Val	Ile	Ala	Leu	Leu 680	Pro	Ala	Asn	Thr	Val 685	Asn	Авр	Asp	
gac Asp	atc Ile 690	gag Glu	atc Ile	tac Tyr	acc Thr	gac Asp 695	gag Glu	acc Thr	cgc Arg	tcg Ser	gaa Glu 700	gtc Val	gcc Ala	ctc Leu	acc Thr	2112
					cag Gln 710											2160
atg Met	cgc	ccg Pro	aac Asn	cgc Arg 725	tgc Cys	ctg Leu	gcg Ala	gac Asp	ttc Phe 730	atc Ile	gcc Ala	ccc Pro	aag Lys	gac Asp 735	acc Thr	2208
ggc Gly	atc Ile	gcc Ala	gat Asp 740	tac Tyr	atc Ile	ggc	ctc Leu	ttc Phe 745	gcg Ala	gtg Val	acg Thr	ggc Gly	ggc Gly 750	atc Ile	gly aaa	2256
atc Ile	gac Asp	aag Lys 755	cgc Arg	gaa Glu	gcc Ala	gcc Ala	ttc Phe 760	gaa Glu	gcc Ala	gac Asp	cac His	gac Asp 765	gac Asp	tac Tyr	agc Ser	2304
gcg	atc Ile 770	atg Met	ctc Leu	aag Lys	gcc Ala	ctg Leu 775	gcc Ala	gac Asp	cgc Arg	ttc Phe	gcc Ala 780	gaa Glu	gcc Ala	ttc Phe	gcc Ala	2352
gag Glu 785	tgc Cys	ctg Leu	cac His	gcc Ala	cgt Arg 790	gtg Val	cgc Arg	cgc Arg	gac Asp	ctg Leu 795	tgg Trp	ggc Gly	tac Tyr	gcg Ala	cag Gln 800	2400
gac Asp	gaa Glu	acg Thr	ctc Leu	gac Asp 805	aac Asn	gac Asp	gcg Ala	ctg Leu	atc Ile 810	cgc Arg	gag Glu	gaa Glu	tac Tyr	cgc Arg 815	ggc Gly	2448
					ggc Gly											2496
cgc Arg	gac Asp	ctg Leu 835	ttc Phe	cgc Arg	gtg Val	ctc Leu	gac Asp 840	gcg Ala	cag Gln	gag Glu	atc Ile	ggc Gly 845	atg Met	aac Asn	ctg Leu	2544
					atg Met											2592
					agc Ser 870											2640
gac Asp	cag Gln	gtg Val	gac Asp	gac Asp 885	atg Met	gcc Ala	gcg Ala	cgc Arg	agc Ser 890	Gly 999	gaa Glu	gac Asp	cgc Arg	cgc Arg 895	aat Asn	2688
					gca Ala				taa							2718

<210> 50 <211> 905

<212> PRT <213> Ralstonia solanacearum

<400> 50

Met Thr Asp His Leu Met Arg Leu Ser Gly Leu Glu Pro Phe Asn Ile

1 5 10 15

Gly Glu Asp Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr
20 25 30

Gly Ser Lys Ala Phe Ala Arg Met Ile Leu Asn Ser Gln Phe Asp Glu 35 40

Ala Leu Ala Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile
50 60

Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Lys Ala Ala Met Val 65 70 75 80

Arg Phe Leu Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Arg Val Pro 85 90 95

Ile Met Ile Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys
100 105 110

Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Glu Gly 115 120 125

Glu Glu Gln Phe Ala His His Ala Lys Leu Ile Lys Arg Tyr Gly Ala 130 135 140

Ala Ala Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Phe 145 150 155 160

Ala Arg Lys Thr Glu Ile Cys Lys Arg Ser Tyr Asp Phe Leu Val Asn 165 170 175

Gln Val Gly Phe Ala Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe 180 185 190

Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Val Asp Phe 195 200 205

Ile Glu Ala Thr Arg Trp Ile Lys Gln Lys Leu Pro His Ala Lys Val 210 215 220

Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Val 225 230 235 240

Val Arg Glu Ala Ile His Thr Val Phe Leu Tyr His Ala Ile Gly Ala 245 250 255

Gly Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Gly Val Tyr Glu 260 265 270

Asn Leu Ala Pro Glu Leu Arg Glu Arg Val Glu Asp Val Val Leu Asn 275 280 285

Arg Arg Pro Asp Ala Thr Asp Arg Leu Leu Glu Ile Ala Asp Arg Tyr 290 295 300

Lys Gly Gly Ala Lys Arg Glu Glu Asn Leu Ala Trp Arg Gln Glu

305 310 315 320

Pro Val Glu Lys Arg Leu Ala His Ala Leu Val His Gly Ile Thr Asp 325 330 335

Tyr Val Val Glu Asp Thr Glu Glu Val Arg Gln Lys Ile Phe Ala Ala 340 345 350

Gly Gly Arg Pro Ile Gln Val Ile Glu Gly Pro Leu Met Asp Gly Met 355 360 365

Asn Ile Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln 370 380

Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile 385 390 395 400

Pro Phe Ile Glu Glu Lys Arg Gln Ile Ala Ala Ala Gly Gly Asp 405 410 415

Val Arg Ser Arg Gly Lys Ile Val Ile Ala Thr Val Lys Gly Asp Val 420 425 430

His Asp Ile Gly Lys Asn Ile Val Thr Val Val Leu Gln Cys Asn Asn 435 440 445

Phe Glu Val Val Asn Met Gly Val Met Val Pro Cys Asn Glu Ile Leu 450 455 460

Ala Lys Ala Lys Val Glu Gly Ala Asp Ile Ile Gly Leu Ser Gly Leu 465 470 475 480

Ile Thr Pro Ser Leu Glu Glu Met Ala Tyr Val Ala Ser Glu Met Gln
485 490 495

Arg Asp Glu Tyr Phe Arg Val Lys Lys Ile Pro Leu Leu Ile Gly Gly 500 505 510

Ala Thr Thr Ser Arg Val His Thr Ala Val Lys Ile Ala Pro Asn Tyr 515 520 525

Glu Gly Pro Val Val Tyr Val Pro Asp Ala Ser Arg Ser Val Ser Val 530 540

Ala Ser Ser Leu Leu Ser Asp Glu Ala Ala Ala Arg Tyr Ile Glu Glu 545 550 555 560

Leu His Ala Asp Tyr Asp Arg Ile Arg Thr Gln His Ala Ser Lys Lys 565 570 575

Ala Met Pro Met Val Ser Leu Ala Ala Ala Arg Ala Asn Lys Thr Arg
580 585 590

Ile Asp Trp Ser Asn Tyr Thr Pro Pro Lys Pro Lys Phe Val Gly Arg
595 600 605

Arg Val Phe Arg Asn Tyr Asp Leu Asn Glu Leu Ala Gln Tyr Ile Asp 610 620

Trp Gly Pro Phe Phe Gln Thr Trp Asp Leu Ala Gly Lys Phe Pro Asp 625 630 635 640

223

Ile Leu Asn Asp Ala Ile Val Gly Glu Ser Ala Arg Arg Val Phe Ser 655

Asp Gly Lys Ser Met Leu Ala Arg Leu Ile Ala Gly Arg Trp Leu Thr 660 665 670

Ala Asn Gly Val Ile Ala Leu Leu Pro Ala Asn Thr Val Asn Asp Asp 685

Asp Ile Glu Ile Tyr Thr Asp Glu Thr Arg Ser Glu Val Ala Leu Thr 690 695 700

Trp Arg Asn Ile Arg Gln Gln Ser Glu Arg Pro Ile Ile Asp Gly Val 705 710 715 720

Met Arg Pro Asn Arg Cys Leu Ala Asp Phe Ile Ala Pro Lys Asp Thr
725 730 735

Gly Ile Ala Asp Tyr Ile Gly Leu Phe Ala Val Thr Gly Gly Ile Gly 740 745 750

Ile Asp Lys Arg Glu Ala Ala Phe Glu Ala Asp His Asp Asp Tyr Ser 755 760 765

Ala Ile Met Leu Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala Phe Ala 770 780

Glu Cys Leu His Ala Arg Val Arg Arg Asp Leu Trp Gly Tyr Ala Gln
785 790 795 800

Asp Glu Thr Leu Asp Asn Asp Ala Leu Ile Arg Glu Glu Tyr Arg Gly 805 810 815

Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Val Lys 820 825 830

Arg Asp Leu Phe Arg Val Leu Asp Ala Gln Glu Ile Gly Met Asn Leu 835 840 845

Thr Glu Ala Leu Ala Met Thr Pro Ala Ala Ser Val Ser Gly Phe Gln 850 855 860

Leu Ser His Pro Asp Ser Thr Tyr Phe Thr Ile Gly Lys Ile Gly Gln 865 870 880

Asp Gln Val Asp Asp Met Ala Ala Arg Ser Gly Glu Asp Arg Arg Asn 885 890 895

Val Glu Arg Ala Leu Ala Pro Asn Leu 900

<210> 51

<211> 3645

<212> DNA

<213> Chlorobium tepidum

<220>

<221> CDS

<222> (1) . . (3642)

<223> RCL00420

<400> 51 gtg ctc gac ggg gcc atg ggc acc atg atc cag agg cat ggc ctc gac Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp 10 gaa cag gac tac cgg ggc gag cgt ttc gct tcg cat gac cat ccg ctg Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu 20 aag ggc aac aac gac ctt ctt gtc atc acc cgg ccc gac atc atc cqt 144 Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg 40 teg ate eac tge gae tte ete gae geg ggt geg gae ate ate gag ace 192 Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr 50 55 tgc acc ttc aac gcc aac ccg atc tcg cag tcg gac tac cag ttg cag Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln 65 70 gac ttg acc cgc gag ctg aac gtg gcg gcg gca aag ata gcc cgc tcg Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser 85 gca gcg gac gag ttc acc gca aag act ccc gac aag ccg cgt ttc gtg Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val 100 gee ggt tee ate gga eeg ace aac aag acg ete teg ete teg eeg gae 384 Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp gtg aac aac ccc ggc ttc cgc gcc gtc acc ttc cag gag atg gtc gat 432 Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp aac tac act gcc cag ctc gaa ggc ttg cac gag ggc ggt gtc gat ctc 480 Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu 155 ttg ctc gtc gag acg gtg ttc gac aca ctg aac tgc aag gcg gcg ctc **528** Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu 170 tac gct atc gag gag tac gcg gtg aaa acc ggc tgg cag gtg ccc gtg 576 Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val 180 185 atg gtc tcc ggc acg gtg gtg gac gcg agc ggc cgc acc ctc tcc ggc Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly 195 200 caa acc acc gag gcg ttc tgg att tcg att tcg cac atg ccg agt ctg 672 Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu 215 ctc tcg gtc ggc ctg aac tgc gca ctc ggc tcc aag cag atg cgc ccc 720 Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro 225 230 ttc atc gag gcg ctc tcg aac atc gcc gaa agc tac gtc agc gtc tat

			_						2	25						
Phe	Ile	Glu	Ala	Leu 245	Ser	Asn	Ile	Ala	Glu 250	Ser	Tyr	Val	Ser	Val 255		•
ccc Pro	aac Asn	gcg Ala	ggc Gly 260	ctg Leu	ccg Pro	aat Asn	gag Glu	ttc Phe 265	Gly	gag Glu	tac Tyr	.gac Asp	gac Asp 270	Ser	ccc Pro	816
gag Glu	tac Tyr	atg Met 275	gcc Ala	gcg Ala	cag Gln	atc Ile	gcg Ala 280	Gly	ttc Phe	gcc Ala	gaa Glu	tca Ser 285	ggc	Phe	gtg Val	864
aac Asn	atc Ile 290	gtc Val	ggc	ggc	tgc Cys	tgc Cys 295	ggc	acc	acg Thr	ccg Pro	acg Thr 300	cac His	atc Ile	cgc Arg	gcc Ala	912
11e 305	Ala	Glu	gcg Ala	Val	Lys 310	Thr	Leu	Pro	Pro	Arg 315	Гув	Arg	Pro	Ala	Asn 320	960
Lys	His	Val	ctg Leu	Arg 325	Leu	Ser	Gly	Leu	Glu 330	Pro	Leu	Val	Val	Asp 335	Glu	1008
Thr	Thr	Gly	ttc Phe 340	Ile	Asn	Val	Gly	Glu 345	Arg	Thr	Asn	Val	Thr 350	Gly	Ser	1056
Arg	Lys	Phe 355	gcc Ala	Arg	Leu	Ile	Lys 360	Glu	Ala	Asn	Tyr	Asp 365	Glu	Ala	Leu	1104
Ser	Ile 370	Ala	cgc Arg	Gln	Gln	Val 375	Glu	Asn	Gly	Ala	Gln 380	Val	Ile	Asp	Val	1152
385	Leu	Asp	gaa Glu	Gly	Met 390	Leu	Asp	Ser	Glu	Lys 395	Val	Ile	Val	Glu	Phe 400	1200
Leu	Asn	Leu	atc Ile	Ala 405	Ser	Glu	Pro	Glu	Ile 410	Ala	Lув	Val	Pro	Val 415	Met	1248
Ile	Asp	Ser	tcg Ser 420	Lys	Trp	Ser	Val	Ile 425	Glu	Asn	Gly	Leu	Arg 430	Сув	Thr	1296
Gln	Gly	Lys 435	agc Ser	Ile	Val	Asn	Ser 440	Ile	Ser	Leu	ГÀВ	Glu 445	Gly	Glu	Glu	1344
Leu	Phe 450	Lys	gag Glu	Arg	Ala	Arg 455	Lys	Ile	Met	Gln	Tyr 460	Gly	Ala	Ala	Ala	1392
Val 465	Val	Met	gcc Ala	Phe	Авр 470	Glu	Gln	Gly	Gln	Ala 475	Asp	Ser	Leu	His	Arg 480	1440
cgc Arg	atc Ile	gag Glu	att Ile	tgc Cys 485	agc Ser	cgc Arg	gcc Ala	Tyr	aaa Lys 490	att Ile	ctc Leu	acc Thr	gaa Glu	gag Glu 495	gtg Val	1488

									_	-20						
ggc	ttc Phe	ccg Pro	ccg Pro 500	gag Glu	gac Asp	atc Ile	atc Ile	ttt Phe 505	gac Asp	ccg Pro	aac Asn	gtg Val	ctg Leu 510	acc Thr	gtg Val	1536
gcc Ala	acc Thr	ggc Gly 515	atc Ile	Asp	gag Glu	cac His	aac Asn 520	aac Asn	tac Tyr	gcg	ctc Leu	gac Asp 525	ttc Phe	atc Ile	gaa Glu	1584
agc Ser	gtg Val 530	cgc Arg	tgg Trp	atc Ile	aag Lys	cag Gln 535	aac Asn	ctg Leu	ccg Pro	cac His	gcg Ala 540	aag Lys	gtc Val	tcc Ser	ggc Gly	1632
ggc Gly 545	atc Ile	agc Ser	aac Asn	gtt Val	tcg Ser 550	ttc Phe	tcc Ser	ttc Phe	cgc Arg	ggc Gly 555	aac Asn	gag Glu	ccg Pro	gtg Val	cgc Arg 560	1680
						ttc Phe										1728
Asp	Met	Gly	11e 580	Val	Asn	gcc Ala	Ala	Gln 585	Leu	Gly	Ile	Tyr	Glu 590	Glu	Ile	1776
gac Asp	ccg Pro	gag Glu 595	ctt Leu	ctt Leu	gtc Val	tat Tyr	gtc Val 600	gag Glu	gac Asp	gtg Val	ctg Leu	ctg Leu 605	aac Asn	cgc Arg	cgc Arg	1824
qaA	Asp 610	Ala	Thr	Glu	Arg	ctc Leu 615	Val	Ala	Phe	Ala	Glu 620	Thr	Ile	Arg	qaA	1872
						gcc Ala										1920
Val	Glu	Glu	Arg	Leu 645	Lys	cac His	Ala	Leu	Val 650	Lys	Gly	Ile	Val	Asp 655	Tyr	1968
						gag Glu										2016
Glu	Val	Ile 675	Glu	Gly	Pro	ctc Leu	Met 680	Asn	Gly	Met	Asn	His 685	Val	Gly	qaA	2064
						atg Met 695										2112
						gta Val										2160
						gac Asp		Ser								2208
gcc	acg	gtg	aag	ggc	gac	gtg	cac	gac	atc	ggc	aag	aac	atc	gtg	tcg	2256

														•		
Ala	Thr	. Val	Lys 740	Gly	Aap	Va]	. His	745	Ile	227 • Gly	/ Lys	neA :	11e		Ser	,
gtg Val	.gtg Val	ctt Leu 755	Ala	tgc Cys	aac Asn	aac Asn	tto Phe	: Asp	gtg Val	l Ile	gac Asp	atc Ile 765	Gly	gto Val	atg Met	2304
atg Met	cca Pro 770	CAB	Asp Asp	aag Lys	att Ile	ctc Leu 775	Glu	gcg	ctg Lev	gca Ala	gaa Glu 780	His	aag Lys	Pro	gac Asp	2352
gtg Val 785	ctc Leu	ggc Gly	ctc Leu	tcc Ser	ggc Gly 790	Leu	atc	acc Thr	ccg	tcg Ser 795	Leu	gaa Glu	gag Glu	atg Met	gcg Ala 800	2400
cac His	gtg Val	gcc Ala	aaa Lys	gag Glu 805	atg Met	gag Glu	cgg Arg	ctc	ggc Gly 810	Met	aac Asn	att Ile	ccg Pro	ctc Leu 815	atc Ile	2448
atc	ggc Gly	ggc Gly	gcg Ala 820	acc Thr	acc Thr	tcg Ser	aag Lys	gtg Val 825	cac His	acg Thr	gcg	gtg Val	aaa Lys 830	ctc Leu	gcg Ala	2496
ccċ Pro	tgc Cys	tac Tyr 835	ccc Pro	agc Ser	ggc Gly	gcg Ala	gta Val 840	gta Val	cac His	gtg Val	ctc Leu	gac Asp 845	gcc Ala	tcg Ser	cgc Arg	2544
agc Ser	gtg Val 850	ccg Pro	gtg Val	gtc Val	agc Ser	aac Asn 855	ctc Leu	tgc Cys	aac Asn	ccc Pro	gcc Ala 860	cag Gln	cgc Arg	gac Asp	agc Ser	2592
tat Tyr 865	atc Ile	gcg Ala	gcg Ala	ctg Leu	aag Lys 870	gat Asp	gag Glu	cag Gln	gag Glu	gcg Ala 875	atg Met	cgc Arg	aag Lys	agc Ser	cac His 880	2640
gcc Ala	gag Glu	cgc Arg	atg Met	gcg Ala 885	gca Ala	aaa Lys	aag Lys	tac Tyr	gtc Val 890	tcg Ser	ctc Leu	gac Asp	gcc Ala	gcc Ala 895	cgc Arg	2688
gac Asp	aac Asn	cgc Arg	ctc Leu 900	acc Thr	att Ile	gac Asp	tgg Trp	gag Glu 905	gcc Ala	gaa Glu	acc Thr	atc Ile	gac Asp 910	aag Lys	ccc Pro	2736
gcc Ala	cag Gln	act Thr 915	ggc	gtc Val	acc Thr	gtg Val	ctg Leu 920	gag Glu	gat Asp	gtc Val	acc Thr	gtc Val 925	ggc	gcg Ala	ctc Leu	2784
Arg	ccg Pro 930	tat Tyr	atc i	gac Asp	tgg Trp	gca Ala 935	mcc Xaa	ttc Phe	ttc Phe	tgg Trp	agc Ser 940	tgg Trp	gag Glu	ctg Leu	cac His	2832
ggc Gly 945	gtc Val	tat Tyr	ccg (Pro (Gln :	att Ile 950	ctg Leu	gag Glu	gat Asp	gaa Glu	aag Lys 955	gtc Val	ggc Gly	gag Glu	gag Glu	gca Ala 960	2880
acc Thr	aaa Lys	ctc Leu	ttc a	aac g Asn 2 965	gac Asp	gcc Ala	acc Thr	gct Ala	ctg Leu 970	ctc Leu	gac Asp	cgg Arg	atc Ile	gac Asp 975	agc Ser	2928
gaa Glu i	aag Lys	Leu :	ctc q Leu (ggc a	atc Ile	aaa Lys	Gly	gtg Val 985	gcg Ala	ggc Gly	atc Ile	ttc Phe	ccg Pro 990	gcc Ala	aac Asn	2976

age ate gge gae gae ate tte gte tat geg gat gae gag ege teg at Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser I 995 1000 1005	le
atc cgc acc gtg ctg cac acc ctg cgc cag caa ggc gaa aag cac gg Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His G 1010 1015 1020	ly
•	ly 940
gtc aac gac tgg atc ggc tgc ttc acc gta acc gcc gga ctc ggc at Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Il 1045 1050 1055	.e
cag aat ttg ctc gac gag ttc aca gca gag aac gac gac tac cac cg Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Ar 1060 1065 1070	g
atc atg aca cag gcg ctc gcc gac cga ctg gcc gaa gcg ttc gca ga Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Gl 1075 1080 1085	u
atg ctg cac gaa aag gtg cgc cgc gaa ctc tgg ggc tac gcg ccc gg Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly 1090 1095 1100	У
gaa atc ctc ggc aac gaa gag ctg atc gcc gaa aag tac cga ggc atc Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile 1105 1110 1115	e 20
cgc ccc gcc ccc ggc tac ccc gcc tgc ccg gat cac acc gaa aag gca Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala 1125 1130 1135	1
atc atc ttc gac ctg ctc aac gct gaa gcg gcc acc ggc gtc acg ctg Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu 1140 1145 1150	ı
acg gaa act ttc gcg atg aac ccc gca gcc tca gtc tgc ggc ctc tac Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr 1155 1160 1165	3504
Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys 1170 1175 1180	
gat cag gtc gaa gac tac gcc aac cgc aaa ggg ctg gaa gta gca gaa Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu 1185	
gcc gag aag tgg ctc gcg ccc tcg ctg aac tac gat cca gcg Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala 1205 1210	3642
taa	3645

<210> 52 <211> 1214

<212> PRT

<213> Chlorobium tepidum

<220>

<221> unsure

<222> 936 .. 936

<223> All occurrences of Xaa indicate any amino acid

<400> 52

Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp
1 10 15

Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu 20 25 30

Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg
35 40 45

Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr 50 55 60

Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln 65 70 75 80

Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser 85 90 95

Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val

Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp 115 120 125

Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp 130 135 140

Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu 145 150 155 160

Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu 165 170 175

Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val 180 185 190

Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly 195 200 205

Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu 210 215 220

Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro 225 230 235 240

Phe Ile Glu Ala Leu Ser Asn Ile Ala Glu Ser Tyr Val Ser Val Tyr
245 250 255

Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp Asp Ser Pro 260 265 270

Glu Tyr Met Ala Ala Gln Ile Ala Gly Phe Ala Glu Ser Gly Phe Val 275 280 285

230 Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Thr His Ile Arg Ala Ile Ala Glu Ala Val Lys Thr Leu Pro Pro Arg Lys Arg Pro Ala Asn 315 Lys His Val Leu Arg Leu Ser Gly Leu Glu Pro Leu Val Val Asp Glu 325 330 Thr Thr Gly Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser 345 Arg Lys Phe Ala Arg Leu Ile Lys Glu Ala Asn Tyr Asp Glu Ala Leu Ser Ile Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile Asp Val Asn Leu Asp Glu Gly Met Leu Asp Ser Glu Lys Val Ile Val Glu Phe 390 Leu Asn Leu Ile Ala Ser Glu Pro Glu Ile Ala Lys Val Pro Val Met 405 Ile Asp Ser Ser Lys Trp Ser Val Ile Glu Asn Gly Leu Arg Cys Thr Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu 440 Leu Phe Lys Glu Arg Ala Arg Lys Ile Met Gln Tyr Gly Ala Ala Ala Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Ser Leu His Arg 470 Arg Ile Glu Ile Cys Ser Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val 490 Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val 505 Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile Glu Ser Val Arg Trp Ile Lys Gln Asn Leu Pro His Ala Lys Val Ser Gly 535 Gly Ile Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu Ala Met His Thr Ala Phe Leu Tyr His Ala Ile His Ala Gly Leu 570 Asp Met Gly Ile Val Asn Ala Ala Gln Leu Gly Ile Tyr Glu Glu Ile Asp Pro Glu Leu Leu Val Tyr Val Glu Asp Val Leu Leu Asn Arg Arg

600

Asp Asp Ala Thr Glu Arg Leu Val Ala Phe Ala Glu Thr Ile Arg Asp

231 610 615 620 Gly Gly Glu Lys Ala Glu Ala Lys Asn Ala Glu Trp Arg Asn Ala Pro 630 635 Val Glu Glu Arg Leu Lys His Ala Leu Val Lys Gly Ile Val Asp Tyr Ile Asp Glu Asp Thr Glu Glu Ala Arg Gln Leu Tyr Pro Ser Pro Leu 665 . Glu Val Ile Glu Gly Pro Leu Met Asn Gly Met Asn His Val Gly Asp Leu Phe Ala Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Arg Ser Val Ala Ala Leu Ile Pro Tyr Ile Glu Glu 710 Glu Lys Ser Lys Asn Cys Asp Thr Ser Ala Lys Ala Lys Val Leu Leu 730 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser 745 Val Val Leu Ala Cys Asn Asn Phe Asp Val Ile Asp Ile Gly Val Met 760 Met Pro Cys Asp Lys Ile Leu Glu Ala Leu Ala Glu His Lys Pro Asp Val Leu Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala His Val Ala Lys Glu Met Glu Arg Leu Gly Met Asn Ile Pro Leu Ile 810 Ile Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Leu Ala Pro Cys Tyr Pro Ser Gly Ala Val Val His Val Leu Asp Ala Ser Arg Ser Val Pro Val Val Ser Asn Leu Cys Asn Pro Ala Gln Arg Asp Ser 855 Tyr Ile Ala Ala Leu Lys Asp Glu Gln Glu Ala Met Arg Lys Ser His 875 Ala Glu Arg Met Ala Ala Lys Lys Tyr Val Ser Leu Asp Ala Ala Arg Asp Asn Arg Leu Thr Ile Asp Trp Glu Ala Glu Thr Ile Asp Lys Pro Ala Gln Thr Gly Val Thr Val Leu Glu Asp Val Thr Val Gly Ala Leu Arg Pro Tyr Ile Asp Trp Ala Xaa Phe Phe Trp Ser Trp Glu Leu His

935

940

232

Gly Val Tyr Pro Gln Ile Leu Glu Asp Glu Lys Val Gly Glu Glu Ala 945 950 955 960

Thr Lys Leu Phe Asn Asp Ala Thr Ala Leu Leu Asp Arg Ile Asp Ser 965 970 975

Glu Lys Leu Gly Ile Lys Gly Val Ala Gly Ile Phe Pro Ala Asn 980 985 990

Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser Ile 995 1000 1005

Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His Gly 1010 1015 1020

Glu Ala Asn Leu Ala Leu Ala Asp Phe Val Ala Pro Arg Glu Ser Gly 1025 1030 1035 1040

Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Ile 1045 1050 1055

Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Arg 1060 1065 1070

Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1075 1080 1085

Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly 1090 1095 1100

Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile 1105 1110 1115 1120

Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala 1125 1130 1135

Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu 1140 1145 1150

Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr 1155 1160 1165

Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys 1170 1175 1180

Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu 1185 1190 1195 1200

Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala

<210> 53

<211> 52

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 53

cccgggatcc gctagcggcg cgccggccgg cccggtgtga aataccgcac ag

<210>	54	
<211>	53	·
<212>	DNA	
<213>	Kûnstliche Sequenz	
		•
<220>	•	
<223>	Beschreibung der künstlichen Sequenz:PCR primer	
		h
<400>		
tctaga	actog agoggoogog googgoottt aaattgaaga ogaaagggoo	tcg 53
<210>		
<211>	 -	
<212>		
<213>	Künstliche Sequenz	
-000-		
<220>	Basalus (human dan 180 at 21 at an an an	
<223>	Beschreibung der künstlichen Sequenz:PCR primer	
-400-	ce .	•
<400>		
gagaco	taga cccggggatc cgctagcggg ctgctaaagg aagcgga	47
<210>	56	**
<211>		
<212>	· · · · · · · · · · · · · · · · · · ·	4
	Künstliche Sequenz	
	Manuellone bedacing	•
<220>	•	•
	Beschreibung der künstlichen Sequenz:PCR primer	
	· · · · · · · · · · · · · · · · · · ·	•
<400>	56	
gagagg	cgcg ccgctagcgt gggcgaagaa ctccagca	38
	•	
<210>	57	
<211>	34 .	
<212>		
<213>	Kûnstliche Sequenz	
<220>		
<223>	Beschreibung der künstlichen Sequenz:PCR primer	
<400>	••	
gagagg	gcgg ccgcgcaaag tcccgcttcg tgaa	34
.010.	F0	
<210>		
<211><212>		
	Künstliche Sequenz	
~~13>	vaneritene sedaens	
<220>		
	Beschreibung der künstlichen Sequenz:PCR primer	
	beschietbung der kunschlenen Sequenz: PCR primer	
<400>	5R	
	graa cractcaaat cootoaagoo acoo	

```
<210> 59
<211> 140
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer
<400> 59
tcgaatttaa atctcgagag gcctgacgtc gggcccggta ccacgcgtca tatgactagt 60
teggacetag ggatategte gacategatg etettetgeg ttaattaaca attgggatee 120
tctagacccg ggatttaaat
<210> 60
<211> 140
<212> DNA
<213> Künstliche Sequenz
<223> Beschreibung der künstlichen Sequenz:PCR primer
gatcatttaa atcccgggtc tagaggatcc caattgttaa ttaacgcaga agagcatcga 60
tgtcgacgat atccctaggt ccgaactagt catatgacgc gtggtaccgg gcccgacgtc 120
aggcctctcg agatttaaat
<210> 61
<211> 33
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer
<400> 61
gagageggee geegateett tttaacceat cae
                                                                   33
<210> 62
<211> 32
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer
<400> 62
aggagcggcc gccatcggca ttttcttttg cg
                                                                   32
<210> 63
<211> 5091
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:Plasmid
gccgcgactg cettegegaa gccttgcccc gcggaaattt cetccaccga gttcgtgcac 60
```

235

acccctatgc caagettett teaccctaaa ttegagagat tggattetta cegtggaaat 120 tettegeaaa aategteece tgategeett tgegacgttg gegteggtge egetggttge 180 gettggettg accgaettga tcageggeeg etcgatttaa atetegagag geetgaegte 240 gggcccggta ccacgcgtca tatgactagt tcggacctag ggatatcgtc gacatcgatg 300 ctcttctgcg ttaattaaca attgggatcc tctagacccg ggatttaaat cgctagcggg 360 ctgctaaagg aagcggaaca cgtagaaagc cagtccgcag aaacggtgct gaccccggat 420 gaatgtcagc tactgggcta tctggacaag ggaaaacgca agcgcaaaga gaaagcaggt 480 agettgeagt gggettacat ggegataget agaetgggeg gttttatgga cageaagega 540 accggaattg ccagctgggg cgccctctgg taaggttggg aagccctgca aagtaaactg 600 gatggettte ttgccgccaa ggatetgatg gegeagggga teaagatetg atcaagagae 660 aggatgagga tegtttegea tgattgaaca agatggattg caegeaggtt etceggeege 720 ttgggtggag aggctattcg gctatgactg ggcacaacag acaatcggct gctctgatgc 780 egeegtgtte eggetgteag egeaggggeg eeeggttett tttgteaaga eegaeetgte 840 cggtgccctg aatgaactgc aggacgaggc agcgcggcta tcgtggctgg ccacgacggg 900 cottectige geagetgige tegacgitgi caetgaageg ggaagggaet ggetgetatt 960 gggcgaagtg ccggggcagg atctcctgtc atctcacctt gctcctgccg agaaagtatc 1020 catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct gcccattcga 1080 ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgtcga 1140 tcaggatgat ctggacgaag agcatcaggg gctcgcgcca gccgaactgt tcgccaggct 1200 caaggegege atgeeegacg gegaggatet egtegtgace catggegatg cetgettgee 1260 gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt 1320 ggcggaccgc tatcaggaca tagcgttggc tacccgtgat attgctgaag agcttggcgg 1380 cgaatggget gaccgettee tegtgettta eggtategee geteeegatt egeagegeat 1440 cgccttctat cgccttcttg acgagttctt ctgagcggga ctctggggtt cgaaatgacc 1500 gaccaagega egeccaacet gecateaega gatttegatt ceaeegeege ettetatgaa 1560 aggttggget teggaategt ttteegggae geeggetgga tgateeteea gegeggggat 1620 ctcatgctgg agttcttcgc ccacgctagc ggcgcccgg ccggcccggt gtgaaatacc 1680 gcacagatgc gtaaggagaa aataccgcat caggcgctct tecgetteet egeteactga 1740 ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 1800 acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca 1860 aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc tccgccccc 1920 tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga caggactata 1980 aagataccag gegttteece etggaagete eetegtgege teteetgtte egaceetgee 2040 gettacegga tacetgteeg cettteteee ttegggaage gtggegettt eteatagete 2100 acgetgtagg tateteagtt eggtgtaggt egttegetee aagetggget gtgtgeaega 2160 accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 2220 ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 2280 gtatgtagge ggtgetacag agttettgaa gtggtggeet aactaegget acaetagaag 2340 gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 2400 ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt tttttgttt gcaagcagca 2460 gattacgcgc agaaaaaaag gatctcaaga agatcctttg atctttcta cggggtctga 2520 cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat 2580 cttcacctag atccttttaa aggccggccg cggccgcgca aagtcccgct tcgtgaaaat 2640 tttcgtgccg cgtgattttc cgccaaaaac tttaacgaac gttcgttata atggtgtcat 2700 gaccttcacg acgaagtact aaaattggcc cgaatcatca gctatggatc tctctgatgt 2760 cgcgctggag tccgacgcgc tcgatgctgc cgtcgattta aaaacggtga tcggattttt 2820 ccgagetete gatacgacgg acgegecage atcacgagae tgggccagtg ccgcgagega 2880 cctagaaact ctcgtggcgg atcttgagga gctggctgac gagctgcgtg ctcggccagc 2940 gccaggagga cgcacagtag tggaggatgc aatcagttgc gcctactgcg gtggcctgat 3000 tecteceegg cetgaceege gaggacggeg egcaaaatat tgeteagatg egtgtegtge 3060 cgcagccagc cgcgagcgcg ccaacaaacg ccacgccgag gagctggagg cggctaggtc 3120 gcaaatggcg ctggaagtgc gtcccccgag cgaaattttg gccatggtcg tcacagagct 3180 ggaagcggca gcgagaatta tcgcgatcgt ggcggtgccc gcaggcatga caaacatcgt 3240 aaatgeegeg tttegtgtge egtggeegee caggaegtgt cagegeegee accaectgca 3300 ccgaatcggc agcagcgtcg cgcgtcgaaa aagcgcacag gcggcaagaa gcgataagct 3360 gcacgaatac ctgaaaaatg ttgaacgccc cgtgagcggt aactcacagg gcgtcggcta 3420 acccccagtc caaacctggg agaaagcgct caaaaatgac tctagcggat tcacgagaca 3480 ttgacacacc ggcctggaaa ttttccgctg atctgttcga cacccatccc gagctcgcgc 3540 tgcgatcacg tggctggacg agcgaagacc gccgcgaatt cctcgctcac ctgggcagag 3600 aaaatttcca gggcagcaag acccgcgact tcgccagcgc ttggatcaaa gacccggaca 3660 cggagaaaca cagccgaagt tataccgagt tggttcaaaa tcgcttgccc ggtgccagta 3720 tgttgctctg acgcacgcgc agcacgcagc cgtgcttgtc ctggacattg atgtgccgag 3780

```
ccaccaggcc ggcgggaaaa tcgagcacgt aaaccccgag gtctacgcga ttttggagcg 3840
ctgggcacgc ctggaaaaag cgccagcttg gatcggcgtg aatccactga gcgggaaatg 3900
ccagctcatc tggctcattg atccggtgta tgccgcagca ggcatgagca gcccgaatat 3960
gegeetgetg getgeaacga eegaggaaat gaeeegegtt tteggegetg aeeaggettt 4020
ttcacatagg ctgagccgtg gccactgcac tctccgacga tcccagccgt accgctggca 4080
tgcccagcac aatcgcgtgg atcgcctagc tgatcttatg gaggttgctc gcatgatctc 4140
aggcacagaa aaacctaaaa aacgctatga gcaggagttt tctagcggac gggcacgtat 4200
cgaagcggca agaaaagcca ctgcggaagc aaaagcactt gccacgcttg aagcaagcct 4260
gccgagcgcc gctgaagcgt ctggagagct gatcgacggc gtccgtgtcc tctggactgc 4320
tecagggegt geogeoogtg atgagaegge ttttegecae getttgaetg tgggatacca 4380
gttaaaagcg gctggtgagc gcctaaaaga caccaagggt catcgagcct acgagcgtgc 4440
ctacaccgtc gctcaggcgg tcggaggagg ccgtgagcct gatctgccgc cggactgtga 4500
ccgccagacg gattggccgc gacgtgtgcg cggctacgtc gctaaaggcc agccagtcgt 4560
ccctgctcgt cagacagaga cgcagagcca gccgaggcga aaagctctgg ccactatggg 4620
aagacgtggc ggtaaaaagg ccgcagaacg ctggaaagac ccaaacagtg agtacgcccg 4680
agcacagcga gaaaaactag ctaagtccag tcaacgacaa gctaggaaag ctaaaggaaa 4740
togottgacc attgcaggtt ggtttatgac tgttgaggga gagactggct cgtggccgac 4800
aatcaatgaa gctatgtctg aatttagcgt gtcacgtcag accgtgaata gagcacttaa 4860
ggtctgcggg cattgaactt ccacgaggac gccgaaagct tcccagtaaa tgtgccatct 4920
cgtaggcaga aaacggttcc cccgtagggt ctctctcttg gcctcctttc taggtcgggc 4980
tgattgctct tgaagctctc taggggggct cacaccatag gcagataacg ttccccaccg 5040
gctcgcctcg taagcgcaca aggactgctc ccaaagatct tcaaagccac t
```

```
<210> 64
<211> 4323
<212> DNA
<213> Kunstliche Sequenz
```

<220>

<223> Beschreibung der künstlichen Sequenz:Plasmid

```
<400> 64
teteteageg tatggttgte geetgagetg tagttgeett categatgaa etgetgtaca 60
ttttgatacg tttttccgtc accgtcaaag attgatttat aatcctctac accgttgatg 120
ttcaaagagc tgtctgatgc tgatacgtta acttgtgcag ttgtcagtgt ttgtttgccg 180
taatgtttac cggagaaatc agtgtagaat aaacggattt ttccgtcaga tgtaaatgtg 240
gctgaacctg accattcttg tgtttggtct tttaggatag aatcatttgc atcgaatttg 300
togotytott taaagacgcg gocagcgttt ttocagctgt caatagaagt ttogocgact 360
ttttgataga acatgtaaat cgatgtgtca tccgcatttt taggatctcc ggctaatgca 420
aagacgatgt ggtagccgtg atagtttgcg acagtgccgt cagcgttttg taatggccag 480
ctgtcccaaa cgtccaggcc ttttgcagaa gagatatttt taattgtgga cgaatcaaat 540
tcagaaactt gatatttttc atttttttgc tgttcaggga tttgcagcat atcatggcgt 600
gtaatatggg aaatgccgta tgtttcctta tatggctttt ggttcgtttc tttcgcaaac 660
gcttgagttg cgcctcctgc cagcagtgcg gtagtaaagg ttaatactgt tgcttgtttt 720
gcaaactttt tgatgttcat cgttcatgtc tcctttttta tgtactgtgt tagcggtctg 780
cttcttccag ccctcctgtt tgaagatggc aagttagtta cgcacaataa aaaaagacct 840
aaaatatgta aggggtgacg ccaaagtata cactttgccc tttacacatt ttaggtcttg 900
cotgetttat cagtaacaaa cocgegegat ttactttteg accteattet attagactet 960
cgtttggatt gcaactggtc tattttcctc ttttgtttga tagaaaatca taaaaggatt 1020
tgcagactac gggcctaaag aactaaaaaa tctatctgtt tcttttcatt ctctgtattt 1080
tttatagttt ctgttgcatg ggcataaagt tgccttttta atcacaattc agaaaatatc 1140
ataatatoto atttoactaa ataatagtga acggcaggta tatgtgatgg gttaaaaagg 1200
ateggeggee getegattta aatetegaga ggeetgaegt egggeeeggt aceaegegte 1260
atatgactag ttcggaccta gggatatcgt cgacatcgat gctcttctgc gttaattaac 1320
aattgggatc ctctagaccc gggatttaaa tcgctagcgg gctgctaaag gaagcggaac 1380
acgtagaaag ccagtccgca gaaacggtgc tgaccccgga tgaatgtcag ctactgggct 1440
atctggacaa gggaaaacgc aagcgcaaag agaaagcagg tagcttgcag tgggcttaca 1500
tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt gccagctggg 1560
gcgccctctg gtaaggttgg gaagccctgc aaagtaaact ggatggcttt cttgccgcca 1620
aggatctgat ggcgcagggg atcaagatct gatcaagaga caggatgagg atcgtttcgc 1680
```

```
237
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 1740
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 1800
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 1860
caggacgagg cagegeget ategtggetg gecaegaegg gegtteettg egeagetgtg 1920
ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag 1980
gateteetgt cateteacet tgeteetgee gagaaagtat ceateatgge tgatgeaatg 2040
cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 2100
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 2160
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac 2220
ggcgaggate tegtegtgae ceatggegat geetgettge egaatateat ggtggaaaat 2280
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 2340
atagegtigg ctaccegiga tatigeigaa gagetiggeg gegaatggge igacegetic 2400
ctcgtgcttt acggtatcgc cgctcccgat tcgcagegca tcgccttcta tcgccttctt 2460
gacgagttet tetgageggg actetggggt tegaaatgae egaceaageg acgeecaace 2520
tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 2580
ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 2640
cccacgctag cggcgcccg gccggcccgg tgtgaaatac cgcacagatg cgtaaggaga 2700
aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 2760
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 2820
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 2880
aaggccgcgt tgctggcgtt tttccatagg ctccgcccc ctgacgagca tcacaaaaat 2940
cgacgeteaa gteagaggtg gegaaaceeg acaggaetat aaagatacea ggegttteec 3000
cetggaaget ceetegtgeg eteteetgtt cegaccetge egettacegg atacetgtee 3060
geetttetee ettegggaag egtggegett teteataget caegetgtag gtateteagt 3120
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac 3180
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 3240
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 3300
gagttettga agtggtggee taactacgge tacactagaa ggacagtatt tggtatetge 3360
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 3420
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaa 3480
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 3540
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 3600
aaggccggcc gcggccgcca tcggcatttt cttttgcgtt tttatttgtt aactgttaat 3660
tgtccttgtt caaggatgct gtctttgaca acagatgttt tcttgccttt gatgttcagc 3720
aggaageteg gegeaaaegt tgattgtttg tetgegtaga atectetgtt tgteatatag 3780
cttgtaatca cgacattgtt tcctttcgct tgaggtacag cgaagtgtga gtaagtaaag 3840
gttacatcgt taggatcaag atccattttt aacacaaggc cagttttgtt cagcggcttg 3900
tatgggccag ttaaagaatt agaaacataa ccaagcatgt aaatatcgtt agacgtaatg 3960
ccgtcaatcg tcatttttga tccgcgggag tcagtgaaca ggtaccattt gccgttcatt 4020
ttaaagacgt tcgcgcgttc aatttcatct gttactgtgt tagatgcaat cagcggtttc 4080
atcacttttt tcagtgtgta atcatcgttt agctcaatca taccgagagc gccgtttgct 4140
aactcagccg tgcgtttttt atcgctttgc agaagttttt gactttcttg acggaagaat 4200
gatgtgcttt tgccatagta tgctttgtta aataaagatt cttcgccttg gtagccatct 4260
tcagttccag tgtttgcttc aaatactaag tatttgtggc ctttatcttc tacgtagtga 4320
gga
<210>
      65
<211>
      35
<212> DNA
<213> PCR Primer
<400> 65
gagagagaga cgcgtcccag tggctgagac gcatc
                                                                      35
<210>
      66
<211>
      34
<212>
      DNA
```

<213> PCR Primer <400> 66 ctctctctgt cgacgaattc aatcttacgg cctg

<400> 74

<210>	67					•	
<211>	38						
<212>	DNA						
<213>		Primer					
<400>	67	PIIMEI					
				h			
cygcat	ccace	gacateatet	tcacctgccc	tegtteeg			38
.210.					•		
<210>	68						
<211>	38						
<212>	DNA	Bandan and		,	•		
<213>		Primer					
<400>	68						
cggaac	gagg	gcaggtgaag	atgatgtcgg	tggtgccg			38
						•	
-210-	C 0						
<210>	69					•	
<211><212>	31						
	DNA	Durimon					
<213> <400>	69	Primer					
				_			
gagact	.cyay	ggaaggtgaa	tcgaatttcg	9 .			31
<210>	70						
<211>	38						
<212>	DNA						
<213>		Primer				•	
<400>	70	FIIMEL					
		G22CCC2CC2	+++++++++++				
greery	13334	gaacgcacga	ttctccaaaa	acaacege			38
<210>	71				•		
<211>	23						
<212>	DNA						
<213>		Primer					
<400>	71						
		ttctccccgg	gac				23
,	1-3-3		340				23
							•
<210>	72						
<211>	22						
<212>	DNA						
<213>		Primer					
400>	72						
		gagttgatca	cc				22
,	J-20	J.: 3 3 0 - 0					4.6
(210>	73						
:211>	18						
	DNA						
		Primer					
400>	73						
		gaageteg					18
- 55		J505					19
:210>	74						
:211>	28						
212>							
		Primer					

ga	gagai	tatc	cct	cagc	9 9 9 (cgtt	gaag			239						28
<2: <2: <2:	10> 11> 12> 13>	75 1266 DNA Lys0	5 C Mut	tante	2											
<22	?1> ?2>	CDS (1).	. (12	266)									٠,'		•,	
<40 gtg Val	gcc	75 : ctg : Lev	gto Val	gta Val	cag Glr	, aaa Lys	tat Tyr	ggc	: ggt ' Gly 10	tco Ser	tcg Sei	ctt Leu	gag Glu	agt Ser	gcg Ala	48
gaa Glu	cgc Arg	att Ile	aga Arg 20	aac Asn	gtc Val	gct Ala	gaa Glu	cgg Arg 25	ato Ile	gtt Val	gcc	acc Thr	aag Lys 30	aag Lys	gct Ala	96
gga	aat 'Asn	gat Asp 35	gtc Val	gtg Val	gtt Val	gtc Val	tgc Cys 40	tcc Ser	gca Ala	atg Met	gga Gly	gac Asp 45	acc Thr	acg Thr	gat Asp	144
gaa Glu	ctt Leu 50	cta Leu	gaa Glu	ctt Leu	gca Ala	gcg Ala 55	gca Ala	gtg Val	aat Asn	ccc Pro	gtt Val 60	ccg Pro	cca Pro	gct Ala	cgt Arg	192
gaa Glu 65	atg Met	gat Asp	atg Met	ctc Leu	ctg Leu 70	act Thr	gct Ala	ggt	gag Glu	cgt Arg 75	att Ile	tct Ser	aac Asn	gct Ala	ctc Leu 80	240
gtc Val	gcc	atg Met	gct Ala	att Ile 85	gag Glu	tcc Ser	ctt Leu	ggc	gca Ala 90	gaa Glu	gcc	caa Gln	tct Ser	ttc Phe 95	acg Thr	288
ggc Gly	tct Ser	cag Gln	gct Ala 100	ggt Gly	gtg Val	ctc Leu	acc Thr	acc Thr 105	gag Glu	cgc Arg	cac His	gga Gly	aac Asn 110	gca Ala	cgc Arg	336
att Ile	gtt Val	gat Asp 115	gtc Val	act Thr	cca Pro	ggt Gly	cgt Arg 120	gtg Val	cgt Arg	gaa Glu	gca Ala	ctc Leu 125	gat Asp	gag Glu	ggc Gly	384
aag Lys	atc Ile 130	tgc Cys	att Ile	gtt Val	gct Ala	ggt Gly 135	ttc Phe	cag Gln	ggt Gly	gtt Val	aat Asn 140	aaa Lys	gaa Glu	acc Thr	cgc Arg	432
gat Asp 145	gtc Val	acc Thr	acg Thr	ttg Leu	ggt Gly 150	cgt Arg	ggt Gly	ggt Gly	tct Ser	gac Asp 155	acc Thr	act Thr	gca Ala	gtt Val	gcg Ala 160	480
ttg Leu	gca Ala	gct Ala	gct Ala	ttg Leu 165	aac Asn	gct Ala	gat Asp	gtg Val	tgt Cys 170	gag Glu	att Ile	tac Tyr	tcg Ser	gac Asp 175	gtt Val	528
gac Asp	ggt Gly	vaı	tat Tyr 180	acc Thr	gct Ala	gac Asp	ccg Pro	cgc Arg 185	atc Ile	gtt Val	cct Pro	aat Asn	gca Ala 190	cag Gln	aag Lys	576
ctg	gaa	aag	ctc	agc	ttc	gaa	gaa	atg	ctg	gaa	ctt	gct	gct	gtt	ggc	624

WO 03/087386 PCT/EP03/04010

240

									- 2	240						
Leu	Glu	Lys 195	Leu	Ser	Phe	Glu	Glu 200		Leu	Glu	Leu	Ala 205		Val	Gly	
tcc Ser	aag Lys 210	Ile	ttg Leu	gtg Val	ctg Leu	cgc Arg 215	agt Ser	gtt Val	gaa Glu	tac Tyr	gct Ala 220	cgt Arg	gca Ala	ttc Phe	aat Asn	672
gtg Val 225	Pro	ctt Leu	cgc Arg	gta Val	cgc Arg 230	tcg Ser	tct Ser	tat Tyr	agt Ser	aat Asn 235	gat Asp	ccc Pro	ggc	act Thr	ttg Leu 240	720
att Ile	gcc Ala	ggc	tct Ser	atg Met 245	gag Glu	gat Asp	att Ile	cct Pro	gtg Val 250	gaa Glu	gaa Glu	gca Ala	gtc Val	ctt Leu 255	acc Thr	768
ggt Gly	gtc Val	gca Ala	acc Thr 260	gac Asp	aag Lys	tcc Ser	gaa Glu	gcc Ala 265	aaa Lys	gta Val	acc Thr	gtt Val	ctg Leu 270	ggt Gly	att Ile	816
tcc Ser	gat Asp	aag Lys 275	cca Pro	ggc Gly	gag Glu	gct Ala	gcg Ala 280	Lys	gtt Val	ttc Phe	cgt Arg	gcg Ala 285	ttg Leu	gct Ala	gat Asp	864
gca Ala	gaa Glu 290	atc Ile	aac Asn	att Ile	gac Asp	atg Met 295	gtt Val	ctg Leu	cag Gln	aac Asn	gtc Val 300	tct Ser	tct Ser	gta Val	gaa Glu	912
gac Asp 305	ggc Gly	acc Thr	acc Thr	gac Asp	atc Ile 310	atc Ile	ttc Phe	acc Thr	tgc Cys	cct Pro 315	cgt Arg	tcc Ser	gac Asp	gly	cgc Arg 320	960
cgc Arg	gcg Ala	atg Met	gag Glu	atc Ile 325	ttg Leu	aag Lys	aag Lys	ctt Leu	cag Gln 330	gtt Val	cag Gln	ggc Gly	aac Asn	tgg Trp 335	acc Thr	100B
aat Asn	gtg Val	ctt Leu	tac Tyr 340	yac Aap	gac Asp	cag Gln	gtc Val	ggc Gly 345	aaa Lys	gtc Val	tcc Ser	ctc Leu	gtg Val 350	ggt Gly	gct Ala	1056
ggc Gly	atg Met	aag Lys 355	tct Ser	cac His	cca Pro	ggt Gly	gtt Val 360	acc Thr	gca Ala	gag Glu	ttc Phe	atg Met 365	gaa Glu	gct Ala	ctg Leu	1104
cgc Arg	gat Asp 370	gtc Val	aac Asn	gtg Val	aac Asn	atc Ile 375	gaa Glu	ttg Leu	att Ile	tcc Ser	acc Thr 380	tct Ser	gag Glu	att Ile	cgt Arg	1152
att Ile 385	tcc Ser	gtg Val	ctg Leu	atc Ile	cgt Arg 390	gaa Glu	gat Asp	gat Asp	ctg Leu	gat Asp 395	gct Ala	gct Ala	gca Ala	cgt Arg	gca Ala 400	1200
ttg Leu	cat His	gag Glu	Gln	ttc Phe 405	cag Gln	ctg Leu	ggc Gly	Gly	gaa Glu 410	gac Asp	gaa Glu	gcc Ala	gtc Val	gtt Val 415	tat Tyr	1248
gca Ala	ggc Gly	acc Thr	gga Gly 420	cgc Arg	taa											1266

241

<212> PRT

<213> LysC Mutante

<400> 76

Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala 1 5 10 15

Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala 20 25 .. 30.

Gly Asn Asp Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp 35 40 45

Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg 50 55 60

Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu 65 70 75 80

Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr 85 90 95

Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
100 105 110

Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
115 120 125

Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg 130 135 140

Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala 145 150 155 160

Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val 165 170 175

Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
180 185 190

Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly 195 200 205

Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn 210 215 220

Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu 225 230 235 240

Ile	Ala	Gly	Ser	Met	Glu	Asp	Ile	Pro	Val	Glu	Glu	Ala	Val	Leu	Thr
				245					250					255	

Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile 260 265 270

Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp 275 280 285

Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu 290 295 300

Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg 305 310 315 320

Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr 325 330 335

Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
340 345 350

Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu 355 360 365

Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg 370 375 380

Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala 385 390 395 400

Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
405 410 415

Ala Gly Thr Gly Arg 420

<210> 77

<211> 5860

<212> DNA

<213> Plasmid

<400> 77

cccggtacca cgcgtcccag tggctgagac gcatccgcta aagccccagg aaccctgtgc 60
agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt 120
aactgtcagc acgtagatcg aaaggtgcac aaaggtggcc ctggtcgtac agaaatatgg 180
cggttcctcg cttgagagtg cggaacgcat tagaaacgtc gctgaacgga tcgttgccac 240

caagaagget ggaaatgatg tegtggttgt etgeteegea atgggagaea ceaeggatg	a 30
acttctagaa cttgcagcgg cagtgaatcc cgttccgcca gctcgtgaaa tggatatgc	
cctgactgct ggtgagcgta tttctaacgc tctcgtcgcc atggctattg agtcccttgg	
cgcagaagec caatetttea egggetetea ggetggtgtg eteaceaceg agegeeaegg	
aaacgcacgc attgttgatg tcactccagg tcgtgtgcgt gaagcactcg atgagggcaa	
gatctgcatt gttgctggtt tccagggtgt taataaagaa acccgcgatg tcaccacgtt	
gggtcgtggt ggttctgaca ccactgcagt tgcgttggca gctgctttga acgctgatgt	
gtgtgagatt tactcggacg ttgacggtgt gtataccgct gacccgcgca tcgttcctaa	
tgcacagaag ctggaaaagc tcagcttcga agaaatgctg gaacttgctg ctgttggctc	
caagattttg gtgctgcgca gtgttgaata cgctcgtgca ttcaatgtgc cacttcgcgt	
acgetegtet tatagtaatg ateceggeae tttgattgee ggetetatgg aggatattee	
tgtggaagaa gcagtcctta ccggtgtcgc aaccgacaag tccgaagcca aagtaaccgt	
tctgggtatt tccgataagc caggcgaggc tgcgaaggtt ttccgtgcgt tggctgatgc	
agaaatcaac attgacatgg ttctgcagaa cgtctcttct gtagaagacg ģcaccaccga	
catcacette acetgeeete gtteegaegg eegeegegeg atggagatet tgaagaaget	1140
tcaggttcag ggcaactgga ccaatgtgct ttacgacgac caggtcggca aagtctccct	1200
cgtgggtget ggcatgaagt ctcacccagg tgttaccgca gagttcatgg aagctctgcg	1260
cgatgtcaac gtgaacatcg aattgatttc cacctctgag attcgtattt ccgtgctgat	1320
ccgtgaagat gatctggatg ctgctgcacg tgcattgcat	1380
cgaagacgaa gccgtcgttt atgcaggcac cggacgctaa agttttaaag gagtagtttt	1440
acaatgacca ccatcgcagt tgttggtgca accggccagg tcggccaggt tatgcgcacc	1500
ettttggaag agcgcaattt cccagctgac actgttcgtt tetttgette cccaegttec	1560
gcaggccgta agattgaatt cgtcgacatc gatgctcttc tgcgttaatt aacaattggg	1620
tcctctaga cccgggattt aaatcgctag cgggctgcta aaggaagcgg aacacgtaga	1680
agccagtcc gcagaaacgg tgctgacccc ggatgaatgt cagctactgg gctatctgga	1740
aagggaaaa cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat	1800
gctagactg ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgccct	1860
tggtaaggt tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct	1920
atggcgcag gggatcaaga tetgatcaag agacaggatg aggategttt egcatgattg	1980
acaagatgg attgcacgca ggttctccgg ccgcttgggt ggagaggcta ttcggctatg	2040
ctgggcaca acagacaatc ggctgctctg atgccgccgt gttccggctg tcagcgcagg	2100

	ggcgcccggt	tctttttgtc	aagaccgacc	tgtccggtg	cctgaatgaa	ctgcaggacg	2160
	aggcagcgcg	g gctatcgtgg	ctggccacga	cgggcgttcc	ttgcgcagct	gtgctcgacg	2220
	ttgtcactga	agcgggaagg	gactggctgc	: tattgggcga	agtgccgggg	caggatetee	2280
	tgtcatctca	ccttgctcct	gccgagaaag	tatccatcat	ggctgatgca	atgcggcggc	2340
	tgcatacgct	tgatccggct	acctgcccat	tcgaccacca	agcgaaacat	cgcatcgage	2400
	gagcacgtac	tcggatggaa	gccggtcttg	tcgatcagga	tgatctggac	gaagagcatc	2460
	aggggctcgc	gccagccgaa	ctgttcgcca	ggctcaaggc	gcgcatgccc	gacggcgagg	2520
	atctcgtcgt	gacccatggc	gatgcctgct	tgccgaatat	catggtggaa	aatggccgct	2580
	tttctggatt	catcgactgt	ggccggctgg	gtgtggcgga	ccgctatcag	gacatagcgt	2640
	tggctacccg	tgatattgct	gaagagettg	gcggcgaatg	ggctgaccgc	ttcctcgtgc	2700
	tttacggtat	cgccgctccc	gattcgcagc	gcatcgcctt	ctatcgcctt	cttgacgagt	2760
	tcttctgagc	gggactctgg	ggttcgaaat	gaccgaccaa	gcgacgccca	acctgccatc	2820
	acgagatttc	gattccaccg	ccgccttcta	tgaaaggttg	ggcttcggaa	tcgttttccg	2880
	ggacgccggc	tggatgatcc	tccagcgcgg	ggatctcatg	ctggagttct	tegeceaege	2940
	tagcggcgcg	ccggccggcc	cggtgtgaaa	taccgcacag	atgcgtaagg	agaaaatacc	3000
	gcatcaggcg	ctcttccgct	tcctcgctca	ctgactcgct	gcgctcggtc	gttcggctgc	3060
	ggcgagcggt	atcagctcac	tcaaaggcgg	taatacggtt	atccacagaa	tcaggggata	3120
	acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	3180
	cgttgctggc	gtttttccat	aggeteegee	cccctgacga	gcatcacaaa	aatcgacgct	3240
	caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	cccctggaa	3300
	gctccctcgt	gcgctctcct	gttccgaccc	tgccgcttac	cggatacctg	teegeettte	3360
	tcccttcggg	aagcgtggcg	ctttctcata	gctcacgctg	taggtatctc	agttcggtgt	3420
	aggtcgttcg	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	3480
	ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	tcgccactgg	3540
1	cagcagccac	tggtaacagg	attagcagag	cgaggtatgt	aggcggtgct	acagagttct	3600
1	tgaagtggtg	gcctaactac	ggctacacta	gaaggacagt	atttggtatc	tgcgctctgc	3660
1	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	3720
•	ctggtagcgg	tggtttttt	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	3780
i	agaagatcc	tttgatcttt	tctacggggt	ctgacgctca	gtggaacgaa	aactcacgtt	3840
ł	aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatectt	ttaaaggccg	3900
9	gccgcggccg	ccatcggcat	tttcttttgc	gtttttattt	gttaactgtt	aattgtcctt	3960

gttcaagga	t gctgtcttt	g acaacagat	g ttttcttgc	tttgatgtt	c agcaggaagc	4020
tcggcgcaa	a cgttgattg	t ttgtctgcg	t agaatcctc	gtttgtcata	a țagcttgtaa	4080
tcacgacat	t gtttccttt	c gcttgaggt	a cagcgaagto	j tgagtaagta	a aaggttacat	4140
cgttaggat	c aagatccat	t tttaacaca	a ggccagtttt	gttcagcgg	ttgtatgggc	4200
cagttaaag	a attagaaac	a taaccaagc	a tgtaaatato	gttagacgt	a atgccgtcaa	4260
tcgtcattt	t tgatccgcgg	g gagtcagtg	a acaggtacca	tttgccgttc	: attttaaaga	4320
cgttcgcgc	g ttcaatttca	a totgttacte	g tgttagatgo	: aatcagcggt	ttcatcactt	4380
ttttcagtg	t gtaatcatco	g tttagctca	a tcataccgag	agcgccgttt	gctaactcag	4440
ccgtgcgtt	t tttatcgctt	tgcagaagti	tttgactttc	ttgacggaag	aatgatgtgc	4500
ttttgccat	a gtatgctttg	, ttaaataaa	attcttcgcc	ttggtagcca	tcttcagttc	4560
cagtgtttg	c ttcaaatact	aagtatttgt	ggcctttatc	ttctacgtag	tgaggatctc	4620
tcagcgtate	gttgtcgcct	gagctgtagt	tgccttcatc	gatgaactgo	tgtacatttt	4680
gatacgttt	tccgtcaccg	r tcaaagattg	g atttataatc	ctctacaccg	ttgatgttca	4740
aagagctgto	: tgatgctgat	acgttaactt	gtgcagttgt	cagtgtttgt	ttgccgtaat	4800
					aatgtggctg	4860
					aatttgtcgc	4920
					ccgacttttt	4980
	gtaaatcgat					5040
	gccgtgatag					5100
	caggcetttt					5160
	tttttcattt					5220
	gccgtatgtt					5280
	tcctgccagc					5340
	gttcatcgtt					5400
	cctgtttgaa					5460
	gtgacgccaa					5520
	aacaaacccg					5580
	ctggtctatt					5640
	ctaaagaact					5700
	tgcatgggca					5760
atctcattt	cactaaataa	tagtgaacgg	caggtatata	taataaatta	aaaaaaatca	5820

geggeegee.	gatttaaat	c ccgagagge	tgacgtcggg	3	·	5860
	-					
<400> 78 cccggtacca	cgcgtcccac	g tggctgagad	: gcatccgcta	aagccccag	g aaccctgtgc	. 60
					ttgagcgggt	120
			•		agaaatatgg	180
cggttcctcg	cttgagagtg	g cggaacgcat	tagaaacgtc	gctgaacgga	tcgttgccac	240
caagaaggct	ggaaatgatg	tcgtggttgt	: ctgctccgca	atgggagaca	ccacggatga	300
acttctagaa	cttgcagcgg	, cagtgaatcc	: cgttccgcca	gctcgtgaaa	tggatatgct	360
cctgactgct	ggtgagcgta	tttctaacgo	tetegtegee	atggctattg	agtcccttgg	420
cgcagaagcc	caatctttca	cgggctctca	ggctggtgtg	ctcaccaccg	agcgccacgg	480
aaacgcacgc	attgttgatg	tcactccagg	tcgtgtgcgt	gaagcactcg	atgagggcaa	540
gatctgcatt	gttgctggtt	tccagggtgt	taataaagaa	acccgcgatg	tcaccacgtt	600
gggtcgtggt	ggttctgaca	ccactgcagt	tgcgttggca	gctgctttga	acgctgatgt	660
gtgtgagatt	tactcggacg	ttgacggtgt	gtataccgct	gacccgcgca	tcgttcctaa	720
tgcacagaag	ctggaaaagc	tcagcttcga	agaaatgctg	gaacttgctg	ctgttggctc	780
caagattttg	gtgctgcgca	gtgttgaata	cgctcgtgca	ttcaatgtgc	cacttcgcgt	840
acgctcgtct	tatagtaatg	atcccggcac	tttgattgcc	ggctctatgg	aggatattcc	900
tgtggaagaa	gcagtcctta	ccggtgtcgc	aaccgacaag	tccgaagcca	aagtaaccgt	960
tctgggtatt	tccgataagc	caggcgaggc	tgcgaaggtt	ttccgtgcgt	tggctgatgc	1020
agaaatcaac	attgacatgg	ttctgcagaa	cgtctcttct	gtagaagacg	gcaccaccga	1080
catcatcttc	acctgccctc	gttccgacgg	ccgccgcgcg	atggagatct	tgaagaagct	1140
tcaggttcag	ggcaactgga	ccaatgtgct	ttacgacgac	caggtcggca	aagtctccct	1200
egtgggtgct	ggcatgaagt	ctcacccagg	tgttaccgca	gagttcatgg	aagctctgcg	1260
cgatgtcaac	gtgaacatcg	aattgatttc	cacctctgag	attcgtattt	ccgtgctgat	1320
ccgtgaagat	gatctggatg	ctgctgcacg	tgcattgcat	gagcagttcc	agctgggcgg	1380
gaagacgaa	gccgtcgttt	atgcaggcac	cggacgctaa	agttttaaag	gagtagtttt	1440
acaatgacca	ccatcgcagt	tgttggtgca	accggccagg	tcggccaggt	tatgcgcacc	1500
ttttggaag	agcgcaattt	cccagctgac	actgttcgtt	tctttgcttc	cccacgttcc	1560
gcaggccgta	agattgaatt	cgtcgacatc	gatgetette	tqcqttaatt	aacaattogo	1620

	•	
atcctctaga cccgggattt aaatcgctag cgg		68
aagccagtcc gcagaaacgg tgctgacccc gga		74
caagggaaaa cgcaagcgca aagagaaagc agg	tagettg cagtgggett acatggegat 1	80
agctagactg ggcggtttta tggacagcaa gcg	aaccgga attgccagct ggggcgccct 10	86
ctggtaaggt tgggaagccc tgcaaagtaa act	ggatggc tttcttgccg ccaaggatct 19	92
gatggcgcag gggatcaaga tctgatcaag aga	caggatg aggatcgttt cgcatgattg 19	98
aacaagatgg attgcacgca ggttctccgg ccg	cttgggt ggagaggcta ttcggctatg 20	04
actgggcaca acagacaatc ggctgctctg atg	ccgccgt gttccggctg tcagcgcagg 21	10
ggcgcccggt tctttttgtc aagaccgacc tgt		160
aggeagegeg getategtgg etggeeaega egg		220
ttgtcactga agcgggaagg gactggctgc tat		8 (
tgtcatctca cettgeteet geegagaaag tate		40
tgcatacget tgatccgget acctgcccat tcga		00
gagcacgtac teggatggaa geeggtettg tega		60
aggggetege gecageegaa etgttegeca gget		20
atctcgtcgt gacccatggc gatgcctgct tgcc		80
tttctggatt catcgactgt ggccggctgg gtgt		40
tggctacccg tgatattgct gaagagcttg gcgg		00
tttacggtat cgccgctccc gattcgcagc gcat		60
tcttctgagc gggactctgg ggttcgaaat gacc		20
acgagattte gattecaccg ccgccttcta tgaa		80
ggacgccggc tggatgatcc tccagcgcgg ggat		40
tageggegeg eeggeeggee eggtgtgaaa tace		00
gcatcaggcg ctcttccgct tcctcgctca ctga		50
ggcgagcggt atcagctcac tcaaaggcgg taata		? 0
acgcaggaaa gaacatgtga gcaaaaggcc agcaa		10
cgttgctggc gtttttccat aggctccgcc cccct		0
caagtcagag gtggcgaaac ccgacaggac tataa		0
getecetegt gegeteteet gtteegacee tgeeg		0
tcccttcggg aagcgtggcg ctttctcata gctca		0
aggicgiteg ciccaagetg ggctgtgtgc acgaa	CCCCC cgttcagccc gaccgctgcg 348	^

WO 03/087386

cettateegg taactategt ettgagteea acceggtaag acaegaetta tegecaetgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 3600 tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 3660 tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 3720 ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 3780 aagaagatee tttgatettt tetaeggggt etgaegetea gtggaaegaa aacteaegtt 3840 aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaaggccg 3900 geogeggeeg coateggeat titetitige gittitatit gitaacigit aatigicett 3960 gttcaaggat gctgtctttg acaacagatg ttttcttgcc tttgatgttc agcaggaagc 4020 tcggcgcaaa cgttgattgt ttgtctgcgt agaatcctct gtttgtcata tagcttgtaa 4080 tcacgacatt gtttcctttc gcttgaggta cagcgaagtg tgagtaagta aaggttacat 4140 cgttaggatc aagatccatt tttaacacaa ggccagtttt gttcagcggc ttgtatgggc 4200 cagttaaaga attagaaaca taaccaagca tgtaaatatc gttagacgta atgccgtcaa 4260 togtoatttt tgatoogogg gagtoagtga acaggtacca tttgcogtto attttaaaga 4320 cgttcgcgcg ttcaatttca tctgttactg tgttagatgc aatcagcggt ttcatcactt 4380 ttttcagtgt gtaatcatcg tttagctcaa tcataccgag agcgccgttt gctaactcag 4440 ccgtgcgttt tttatcgctt tgcagaagtt tttgactttc ttgacggaag aatgatgtgc 4500 ttttgccata gtatgctttg ttaaataaag attcttcgcc ttggtagcca tcttcagttc 4560 cagigiting ticaaatact aagtatiigi ggcctttatc tictacgiag igaggatcic 4620 tcagcgtatg gttgtcgcct gagctgtagt tgccttcatc gatgaactgc tgtacatttt 4680 gatacgtttt tccgtcaccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740 aagagetgte tgatgetgat aegttaaett gtgeagttgt eagtgtttgt ttgeegtaat 4800 gtttaccgga gaaatcagtg tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860 aacctgacca ttcttgtgtt tggtctttta ggatagaatc atttgcatcg aatttgtcgc 4920 tgtctttaaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980 gatagaacat gtaaatcgat gtgtcatccg catttttagg atctccggct aatgcaaaga 5040 cgatgtggta gccgtgatag tttgcgacag tgccgtcagc gttttgtaat ggccagctgt 5100 cccaaacgtc caggcctttt gcagaagaga tattttaat tgtggacgaa tcaaattcag 5160 aaacttgata tttttcattt ttttgctgtt cagggatttg cagcatatca tggcgtgtaa 5220 tatgggaaat gccgtatgtt tccttatatg gcttttggtt cgtttctttc gcaaacgctt 5280 gagttgcgcc tcctgccagc agtgcggtag taaaggttaa tactgttgct tgttttgcaa 5340

actttttgat	gttcatcgtt	catgtctcct	ttttatgta	ctgtgttagc	ggtctgcttc	5400
ttccagccct	cctgtttgaa	gatggcaagt	tagttacgca	caataaaaaa	agacctaaaa	5460
tatgtaaggg	gtgacgccaa	agtatacact	ttgcccttta	cacattttag	gtcttgcctg	5520
ctttatcagt	aacaaacccg	cgcgatttac	ttttcgacct	cattctatta	gactctcgtt	5580
tggattgcaa	ctggtctatt	ttcctcttt	gtttgataga	aaatcataaạ	aggatttgca	5640
gactacgggc	ctaaagaact	aaaaaatcta	tctgtttctt	ttcattctct	gtattttta	5700
tagtttctgt	tgcatgggca	taaagttgcc	ttttaatca	caattcagaa	aatatcataa	5760
tatctcattt	cactaaataa	tagtgaacgg	caggtatatg	tgatgggtta	aaaaggatcg	5820
gcggccgctc	gatttaaatc	tcgagaggcc	tgacgtcggg			5860
<400> 79	smid					
	gtgaatcgaa					60
	accetttttg				****	120
	cgaccgcgat				•	180
	accgggcaga					240
	gccatccacc					300
	cacccgcgtg					360
	gctggacgac					420
cccggcccga	catcgtccgc	tcggtgcacg	aggagtactt	cgcggccggc	gtcgactgcg	480
tcgagaccaa	caccttcggc	gccaaccact	ccgccctggg	cgagtacgac	atccccgage	540
gcgtccacga	actgtccgag	gccggcgccc	gcgtcgcccg	cgaggtcgcc	gacgagttcg	600
gcgcccgcga	cggccggcag	cgctgggtgc	tgggctccat	gggccccggc	accaagctcc	660
ccaccctcgg	ccacgccccg	tacaccgtcc	tgcgcgacgc	ctaccagcgc	aacgccgagg	720
gactggtcgc	gggcggcgcg	gacgcactgc	tggtggagac	cacgcaggac	ctgctccaga	780
ccaaggcctc	ggtgctcggc	gcccggcgcg	ccctggacgt	cctcggcctc	gacctgccgc	840
catcgtgtc	cgtcaccgtc	gagaccaccg	gcaccatgct	gctcggctcg	gagatcggcg	900
ccgcgctcac	cgcgctggaa	ccgctcggca	tcgacatgat	cggcctgaac	tgcgccaccg	960
gccccgccga	gatgagcgag	cacctgcgct	acctcgcccg	gcactcccgc	atcccgctga	1020
cctgcatgcc	caacgccggt	ctgcccgtcc	tcggcaagga	cggcgcccac	tacccgctga	.1080

ccgcgcccga gctggccgac gcacacgaga ccttcgtgcg cgagtacggc ctgtccctgg 1140

teggegget	g ctgcggcac	c acgcccgage	acctgcgcca	ggtcgtcgag	cgggteeggg	1200
acaccgccc	c caccgcacg	gacccgcgc	ccgageccgg	cgccgcctcg	ctctaccaga	1260
ccgtgccctt	ccgccaggad	e acctcctaco	: tggccatcgg	cgagcgcacc	aacgccaacg	1320
ggtccaagaa	gttccgcgag	g gecatgetgg	acggccgctg	ggacgactgc	gtcgagatgg	1380
cccgcgacca	gatccgcgaa	a ggcgcgcaca	tgctcgacct	ctgcgtcgac	tacgtcggcc	1440
gggacggcgt	cgccgacat	gaggaactgg	ccggccggtt	cgccaccgcc	tccacgctgc	1500
cgatcgtcct	cgactccac	gaggtcgacg	tcatccgggc	cggcctggag	aagctcggcg	1560
gccgcgc gg t	gatcaactcg	gtcaactacg	aggacggcgc	cggccccgag	tcccggttcg	1620
cccgcgtcac	gaagetegee	: cgggagcacg	gcgccgcgct	gatcgcgctg	accatcgacg	1680
aggtgggaca	ggcccgcacc	: gccgagaaga	aggtcgagat	cgccgaacgg	ctcatcgacg	1740
acctcaccgg	caactggggc	: atccacgagt	ccgacatcct	cgtcgactgc	ctgaccttca	1800
ccatctgcac	cggccaggag	gagtcccgca	_aggacggcct	ggccaccatc	gagggcatcc	1860
gggaactcaa	gcggcgccac	ccggacgtgc	agaccacgct	cggcctgtcg	aacatctcct	1920
tcggcctcaa	cccggccgcc	cgcatcctgc	tcaactccgt	cttcctcgac	gaatgcgtca	1980
aggccggcct	ggactcggcc	atcgtgcacg	cgagcaagat	cctgccgatc	gcccgcttcg	2040
acgaggagca	ggtcaccacc	gccctcgact	tgatctacga	ccgccgccgc	gagggctacg	2100
accccctgca	aaagctcatg	cagctcttcg	agggcgccac	cgccaagtcg	ctgaaggcct	2160
ccaaggccga	ggaactggcc	gccctcccgc	tggaggagcg	cctcaagcgc	cgcatcatcg	2220
acggcgagaa	gaacggcctc	gaacaggacc	tcgacgaggc	cctccgggag	cgcccggccc	2280
tcgagatcgt	caacgacacc	ctgctcgacg	gtatgaaggt	cgtcggcgag	ctgttcggct	2340
ccggccagat	gcagctgccg	ttcgtgctcc	agtccgccga	ggtcatgaag	accgcggtgg	2400
cccacctgga	gccgcacatg	gagaagaccg	acgacgacgg	caagggcacg	atcgtgctgg	2460
caccgtccg	cggcgacgtc	cacgacatcg	gcaagaacct	cgtcgacatc	atcctgtcca	2520
acaacggcta	caacgtcgtc	aacctcggca	tcaagcagcc	cgtctccgcg	atcctggaag	2580
ggccgacga	gcaccgggcc	gacgtcatcg	gcatgtccgg	cctcctcgtc	aagtccacgg	2640
gatcatgaa	ggagaacctg	gaggagctga	accagcgcaa	gctggccgcc	gactacccgg	2700
cateetegg	cggcgccgcc	ctcaccaggg	cctacgtcga	acaggacctg	cacgagatet	2760
ıcgacggcga	ggtccgctac	gcccgcgacg	ccttcgaggg	cctgcgcctc	atggacgccc	2820
catcggcat	caagcgcggc	gtgcccggcg	ccaagetgcc	ggagctgaag	cagcgccggg	2880
gcgggccgc	caccgtcgag	atcgacgagc	gccccgagga	aggccacgtc	cgctccgacg	2940
cgccaccga	caacccggtc	ccgaccccgc	ccttccgcgg	cacccgcgtc	gtcaagggca	3000

tccagctcaa	a ggagtacgc	c teetggeteg	acgaggggg	cctcttcaag	ggccagtggg	3060
gcctcaagca	ggcccgcac	c ggcgagggad	cctcctacga	ggaactggto	gagtccgagg	3120
gccggccgcg	getgegegge	ctgctcgacc	ggctccagac	ggacaacctt	ttggaggcgg	3180
ccgtggtcta	cggctactto	ccctgcgtct	ccaaggacga	cgacctgato	gtcctcgacg	3240
acgacggcaa	cgaacgcac	cgcttcacct	tcccccgcca	gegeegege	cggcgcctgt	3300
gcctggccga	cttcttccg	ccggaggagt	ccggcgagac	cgacgtggtc	ggcttccagg	3360
tcgtcaccgt	cggeteeege	atcggcgagg	agacggcccg	catgttcgag	gccaacgcct	3420
accgcgacta	tctcgagctg	cacggcctgt	ccgtgcagct	cgccgaggcc	ctcgccgagt	3480
actggcacgc	gegegtgege	tcggaactcg	gcttcgccgg	ggaggacccg	gccgagatgg	3540
aggacatgtt	cgccctgaag	taccggggtg	cccgcttctc	cctcggctac	ggcgcctgcc	3600
ccgacctgga	ggaccgcgcc	: aagatcgccg	ccctgctgga	gcccgagcgc	atcggcgtcc	3660
acctatccga	ggagttccag	ctccaccccg	agcagtccac	cgacgccatc	gtcatccacc	3720
acccggaggc	caagtacttc	aacgcccgct	gagggatatc	gtcgacatcg	atgctcttct	3780
gcgttaatta	acaattggga	tcctctagac	ccgggattta	aatcgctagc	gggctgctaa	3840
aggaagcgga	acacgtagaa	agccagtccg	cagaaacggt	gctgaccccg	gatgaatgtc	3900
agctactggg	ctatctggac	aagggaaaac	gcaagcgcaa	agagaaagca	ggtagcttgc	3960
agtgggctta	catggcgata	gctagactgg	gcggttttat	ggacagcaag	cgaaccggaa	4020
ttgccagctg	gggcgccctc	tggtaaggtt	gggaagccct	gcaaagtaaa	ctggatggct	4080
ttcttgccgc	caaggatctg	atggcgcagg	ggatcaagat	ctgatcaaga	gacaggatga	4140
ggatcgtttc	gcatgattga	acaagatgga	ttgcacgcag	gttctccggc	cgcttgggtg	4200
gagaggctat	tcggctatga	ctgggcacaa	cagacaatcg	gctgctctga	tgccgccgtg	4260
ttccggctgt	cagcgcaggg	gcgcccggtt	ctttttgtca	agaccgacct	gtccggtgcc	4320
ctgaatgaac	tgcaggacga	ggcagcgcgg	ctatcgtggc	tggccacgac	gggcgttcct	4380
gcgcagctg	tgctcgacgt	tgtcactgaa	gcgggaaggg	actggctgct	attgggcgaa	4440
gtgccggggc	aggateteet	gtcatctcac	cttgctcctg	ccgagaaagt	atccatcatg	4500
gctgatgcaa	tgcggcggct	gcatacgctt	gatccggcta	cctgcccatt	cgaccaccaa	4560
gcgaaacatc	gcatcgagcg	agcacgtact	cggatggaag	ccggtcttgt	cgatcaggat	4620
gatctggacg	aagagcatca	ggggctcgcg	ccagccgaac	tgttcgccag	gctcaaggcg	4680
gcatgcccg	acggcgagga	tctcgtcgtg	acccatggcg	atgcctgctt	gccgaatatc	4740
tggtggaaa	atggccgctt	ttctggattc	atcgactgtg	gccggctggg	tgtggcggac	4800
gctatcagg	acatagcgtt	ggctacccgt	gatattqctq	aaqaqcttqq	cqqcqaatqq	4860

gctgaccgc	t teetegtget	ttacggtate	c geegeteeeg	attcgcagcg	catcgccttc	492
tategeett	c ttgacgagtt	cttctgagc	g ggactctggg	gttcgaaatg	accgaccaag	498
cgacgccca	a cctgccatca	a cgagatttc	g attccaccgo	cgccttctat	gaaaggttgg	5040
gcttcggaa	t cgttttccgg	gacgccggct	ggatgatcct	ccagcgcggg	gatctcatgc	510
tggagttcti	cgcccacgct	ageggegeg	cggccggcc	ggtgtgaaat	accgcacaga	5160
tgcgtaagga	a gaaaataccg	, catcaggcg	tetteegett	cctcgctcac	tgactcgctg	5220
cgctcggtcg	g tteggetgeg	gcgagcggta	tcagctcact	: caaaggcggt	aatacggtta	5280
tccacagaat	: caggggataa	cgcaggaaag	, aacatgtgag	, caaaaggcca	gcaaaaggcc	5340
aggaaccgta	aaaaggccgc	gttgctggcg	, tttttccata	ggctccgccc	ccctgacgag	5400
catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	: cgacaggact	ataaagatac	5460
caggcgtttc	ccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	5520
ggatacctgt	cegeetttet	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	5580
aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	5640
gttcagcccg	accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	5700
cacgacttat	cgccactggc	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	5760
ggcggtgcta	cagagttctt	gaagtggtgg	cctaactacg	gctacactag	aaggacagta	5820
tttggtatct	gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	5880
tccggcaaac	aaaccaccgc	tggtagcggt	ggttttttg	tttgcaagca	gcagattacg	5940
cgcagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	6000
tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	6060
tagatccttt	taaaggccgg	ccgcggccgc	gcaaagtccc	gcttcgtgaa	aattttcgtg	6120
ccgcgtgatt	ttccgccaaa	aactttaacg	aacgttcgtt	ataatggtgt	catgacette	6180
acgacgaagt	actaaaattg	gcccgaatca	tcagctatgg	atctctctga	tgtcgcgctg	6240
gagtccgacg	cgctcgatgc	tgccgtcgat	ttaaaaacgg	tgatcggatt	tttccgagct	6300
ctcgatacga	cggacgcgcc	agcatcacga	gactgggcca	gtgccgcgag	cgacctagaa	6360
actctcgtgg	cggatcttga	ggagctggct	gacgagctgc	gtgctcggcc	agcgccagga	6420
ggacgcacag	tagtggagga	tgcaatcagt	tgcgcctact	gcggtggcct	gattcctccc	6480
cggcctgacc	cgcgaggacg	gcgcgcaaaa	tattgctcag	atgcgtgtcg	tgccgcagcc	6540
agccgcgagc	gcgccaacaa	acgccacgcc	gaggagctgg	aggcggctag	gtcgcaaatg	6600
gcgctggaag	tgcgtccccc	gagcgaaatt	ttggccatgg	tcgtcacaga	gctggaagcg	6660
gcagcgagaa	ttatcgcgat	cgtggcggtg	cccgcaggca	tgacaaacat	cqtaaatqcc	6720

gegittegig igeegiggee geecaggaeg igicagegee geeaceacei geacegaate 6780 ggcagcagcg tcgcgcgtcg aaaaagcgca caggcggcaa gaagcgataa ģctgcacgaa 6840 tacctgaaaa atgttgaacg ccccgtgagc ggtaactcac agggcgtcgg ctaaccccca 6900 gtccaaacct gggagaaagc gctcaaaaat gactctagcg gattcacgag acattgacac 6960 accggcctgg aaatttteeg ctgatetgtt egacacecat ceegageteg egetgegate 7020 acgtggctgg acgagcgaag accgccgcga attcctcgct cacctgggca gagaaaattt 7080 ccagggcagc aagacccgcg acttcgccag cgcttggatc aaagacccgg acacggagaa 7140 acacagooga agttatacog agttggttca aaatogottg cooggtgcca gtatgttgct 7200 ctgacgcacg cgcagcacgc agccgtgctt gtcctggaca ttgatgtgcc gagccaccag 7260 gccggcggga aaatcgagca cgtaaacccc gaggtctacg cgattttgga gcgctgggca 7320 cgcctggaaa aagcgccagc ttggatcggc gtgaatccac tgagcgggaa atgccagctc 7380 atctggctca ttgatccggt gtatgccgca gcaggcatga gcagcccgaa tatgcgcctg 7440 ctggctgcaa cgaccgagga aatgacccgc gttttcggcg ctgaccaggc tttttcacat 7500 aggetgagee gtggccactg cacteteega egateeeage egtacegetg geatgeeeag 7560 cacaatcgcg tggatcgcct agctgatctt atggaggttg ctcgcatgat ctcaggcaca 7620 gaaaaaccta aaaaacgcta tgagcaggag ttttctagcg gacgggcacg tatcgaagcg 7680 gcaagaaaag ccactgcgga agcaaaagca cttgccacgc ttgaagcaag cctgccgagc 7740 gccgctgaag cgtctggaga gctgatcgac ggcgtccgtg tcctctggac tgctccaggg 7800 cgtgccgccc gtgatgagac ggcttttcgc cacgctttga ctgtgggata ccagttaaaa 7860 gcggctggtg agcgcctaaa agacaccaag ggtcatcgag cctacgagcg tgcctacacc 7920 gtcgctcagg cggtcggagg aggccgtgag cctgatctgc cgccggactg tgaccgccag 7980 acggattggc cgcgacgtgt gcgcggctac gtcgctaaag gccagccagt cgtccctgct 8040 cgtcagacag agacgcagag ccagccgagg cgaaaagctc tggccactat gggaagacgt 8100 ggcggtaaaa aggccgcaga acgctggaaa gacccaaaca gtgagtacgc ccgagcacag 8160 cgagaaaaac tagctaagtc cagtcaacga caagctagga aagctaaagg aaatcgcttg 8220 accattgcag gttggtttat gactgttgag ggagagactg gctcgtggcc gacaatcaat 8280 gaagctatgt ctgaatttag cgtgtcacgt cagaccgtga atagagcact taaggtctgc 8340 gggcattgaa cttccacgag gacgccgaaa gcttcccagt aaatgtgcca tctcgtaggc 8400 agaaaacggt tcccccgtag ggtctctctc ttggcctcct ttctaggtcg ggctgattgc 8460 tettgaaget etetaggggg geteacacea taggeagata aegtteecea eeggetegee 8520 togtaagogo acaaggactg ctoccaaaga tottcaaago cactgoogog actgoottog 8580

WO 03/087386

PCT/EP03/04010

254

ttgatcagcg	gccgctcgat	ttaaatc				8787
cccctgatcg	cccttgcgac	gttggcgtcg	gtgccgctgg	ttgcgcttgg	cttgaccgac	8760
tctttcaccc	taaattcgag	agattggatt	cttaccgtgg	aaattcttcg	caaaaatcgt	8700
cgaagecetg	ccccgcggaa	atttcctcca	ccgagttcgt	gcacacccct	atgccaagct	8640

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
☐ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
FADED TEXT OR DRAWING
D BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

IMAGES ARE BEST AVAILABLE COPY.

OTHER:

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.